# Optimal Impulsive Control of Piecewise Deterministic Markov Processes

F. Dufour
Institut Polytechnique de Bordeaux
INRIA Bordeaux Sud Ouest, Team: CQFD
IMB Institut de Mathématiques de Bordeaux, Université de Bordeaux, France
200 Avenue de la Vieille Tour, 33405 Talence cedex. France
e-mail: dufour@math.u-bordeaux1.fr

M. Horiguchi
Department of Mathematics and Physics, Faculty of Science
Kanagawa University
2946 Tsuchiya, Hiratsuka-shi, Kanagawa Prefecture, 259-1293, Japan.
e-mail: horiguchi@kanagawa-u.ac.jp

A. B. Piunovskiy
Department of Mathematical Sciences
University of Liverpool
L69 7ZL, Liverpool, United Kingdom.
e-mail: piunov@liv.ac.uk

March 30, 2016

**Abstract**

In this paper, we study the infinite-horizon expected discounted continuous-time optimal control problem for Piecewise Deterministic Markov Processes (PDMPs) with both impulsive and gradual (also called continuous) controls. The set of admissible control strategies is supposed to be formed by policies possibly randomized and depending on the past-history of the process. We assume that the gradual control acts on the jump intensity and on the transition measure, but not on the flow. The so-called Hamilton-Jacobi-Bellman (HJB) equation associated to this optimization problem is analyzed. We provide sufficient conditions for the existence of a solution to the HJB equation and show that the solution is in fact unique and coincides with the value function of the control problem. Moreover, the existence of an optimal control strategy is proven having the property to be stationary and non-randomized.

**Keywords:** optimal control, piecewise deterministic Markov process, impulsive control, discounted cost.

**AMS 2000 subject classification:** Primary 90C40; Secondary 60J25.

# 1 Introduction

The objective of this paper is to study the infinite-horizon expected discounted continuous-time optimal control problem for Piecewise Deterministic Markov Processes (PDMPs) with both impulsive and gradual (continuous) controls. PDMPs have been introduced in the literature by M.H.A. Davis [6] as a general class of Markov processes involving deterministic motion punctuated by random jumps. The motion of the PDMP $\{X(t)\}$ depends on three local characteristics, namely the flow $\phi$, the jump rate $\lambda$ and the transition measure $Q$, which specifies the post-jump location. Starting from $x$ the motion of the process follows the flow $\phi(x,t)$ until the first jump time $T_1$ which occurs either spontaneously in a Poisson-like fashion with rate $\lambda(\phi(x,t))$ or when the flow $\phi(x,t)$ hits the boundary of the state space. In either case the location of the process at the jump time $T_1$: $X(T_1) = Z_1$ is selected according to the transition measure $Q(\cdot|\phi(x,T_1))$. Starting from $Z_1$, the next sojourn time $T_2 - T_1$ and post-jump location $X(T_2) = Z_2$ are selected. This gives a piecewise deterministic trajectory for $\{X(t)\}$ with jump times $\{T_k\}$ and post-jump locations $\{Z_k\}$ which follows the flow $\phi$ between two jumps. A suitable choice of the state space and the local characteristics $\phi$, $\lambda$, and $Q$ provides stochastic models covering a great number of problems of operations research [6].

In this work, we study both types of control as described by M.H.A. Davis in his book [6]: gradual control (also called continuous control in the literature) acting continuously in time on the jump intensity $\lambda$ and on the transition measure $Q$, but not on the flow $\phi$, and impulsive control, used to describe control actions that move the process to a new point of the state space at some specific times. The goal is to minimize the infinite-horizon total expected discounted cost, which is composed of a running cost and a cost associated to the jumps of the process, added to the total cost each time the PDMP jumps. The set of admissible control strategies is assumed to be formed by policies possibly randomized and depending on the past-history of the process. We also allow interventions to occur even at natural jump epochs and at time $t = 0$.

A natural technique to solve the optimization problem under study is to characterize the value function as a solution to the Hamilton-Jacobi-Bellman (HJB) equation. Our results follow this approach. In Section 5, we provide sufficient conditions for the existence of a solution to the integro-differential HJB equation in Theorem 5.5 and show that the solution of this optimality equation is in fact unique and coincides with the value function of the optimization problem under consideration. Moreover, the existence of an optimal control strategy is proven having the property to be stationary and non-randomized in Proposition 5.6.

If the impulsive controls are absent then we have the classical PDMP. In this context, let us mention the work [4] where the past dependent gradual control was considered. If the flow is constant then the model transforms into the standard continuous-time Markov decision process. Both the gradual and impulsive controls were studied in [8, 9] where more references can be found.

Gradual and impulsive controls have been extensively studied in the framework of PDMPs, see for example the book [6], the recent work [11] and the references therein. When compared with the PDMP literature, let us highlight [5, 6, 7] as the closest references to our work. It is important to emphasize that in [6], the author chose to study separately gradual and impulse control problems in order to get stronger results in each of these frameworks (see chapter 4, for continuous control and chapter 5, for impulsive control). In [5, 6, 7], the gradual control strategy was chosen among the set of piecewise open loop policies, that is, stochastic kernels or measurable functions that depend only on the last jump time and post jump location (only measurable functions were considered in [5]). Time interval up to the next intervention is non randomized. For a more detailed description of this class of control, the reader may consult sections 44 and 46 in [6]. Moreover, it was postulated that for any control strategy the process is non explosive. In [6, section 46] and [7], the parameters of the model were roughly speaking Lipschitz continuous and the sets of feasible actions were state independent. It is shown in [6, section 46] that a weak form of the Bellman equation gives a necessary and sufficient condition for optimality in the context of gradual control. Note also that the authors of [6, 7] considered the case of the controlled flow. The definition of interventions in [5] were very specific and the cost functions were assumed to be positive. In [6, section 54], it is shown in the context where only impulsive controls are allowed that the value function is the unique fixed point of the operator associated to the *implicit optimal stopping problem* and assuming that the parameters of the model are loosely speaking Lipschitz continuous, it is proved that the value function is the unique solution of a set of quasi-variational inequalities. It should be emphasized that we consider a broaden class of control strategies (possibly random and depending on the past-history of the process and taking values in the state-dependent action spaces), instead of the open loop policies for the gradual control, deterministic time interval up to the next intervention and fixed action sets as in [6, 7]. Besides, we impose the standard continuity-compactness conditions which are needed for the proof of the solvability of the HJB equation. Another main novelty is that we provide sufficient conditions based on the three local characteristics of the process $\phi$, $\lambda$, $Q$ to guarantee the process is non explosive under any control strategy with finite cost. As far as the authors are aware of, this is the first time that this kind of result is presented in the literature for discounted control problems of PDMPs with impulsive and gradual control considering the broaden class of controls mentioned above. Let us also mention [15], where the flow was supposed not to drive the process outside the state space $X$, and jumps were not compulsory when the process hits the boundary. Note that in the context where both gradual and impulsive controls are allowed, it is proved in [5, 7, 15] that the Bellman function (the infimum of the optimality criterion) satisfies the optimality equation in the integral or differential form. After that, verification theorems provide necessary and sufficient conditions for the optimality of a control strategy. The uniqueness of the solution to the optimality equation was not investigated. On the opposite, in the current paper, similarly to [4], we firstly introduce the optimality equation, prove its solvability, and after that show that the Bellman function is its unique solution in the class of bounded lower semi-continuous functions. Note also that in the current article, simultaneous sequences of impulses are not allowed but, as mentioned in Remark 2.2, this condition is not really a restriction.

We describe the model in Section 2 and formulate the optimization problem and assumptions in Section 3. Preliminary results are given in Section 4 and the Dynamic Programming approach is developed in Section 5 containing our main results. In the last section, we describe possible extensions of our present work.

# 2 The controlled PDMP

The main goal of this section is to introduce the notation, the parameters defining the model, and to present the construction of the controlled process. In particular a measurable space $(\Omega, \mathcal{F})$ consisting of the canonical sample paths of the multivariate point process $(\Theta_n, X_n, A_n^i, \overline{X}_n)$ is introduced. Having defined the class of admissible strategies, we show the existence of a probability measure $\mathbb{P}_{x_0}^u$ with respect to which the controlled process $(\Theta_n, X_n, A_n^i, \overline{X}_n)$ has the required conditional distributions.

The following notation will be used in this paper: $\mathbb{N}$ is the set of natural numbers including 0, $\mathbb{N}^* = \mathbb{N} - \{0\}$, $\mathbb{R}$ denotes the set of real numbers, $\mathbb{R}_+$ the set of non-negative real numbers, $\mathbb{R}_+^* = \mathbb{R}_+ - \{0\}$, $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{+\infty\}$ and $\overline{\mathbb{R}}_+^* = \mathbb{R}_+^* \cup \{+\infty\}$. The term *measure* will always refer to a countably additive, $\mathbb{R}_+$-valued set function. Let $X$ be a Borel space (i.e. a measurable subset of a complete and separable metric space) and denote by $\mathcal{B}(X)$ its associated Borel $\sigma$-algebra.

For any set $A$, $I_A$ denotes the indicator function of the set $A$. The set of measures defined on $(X, \mathcal{B}(X))$ is denoted by $\mathcal{M}(X)$, and $\mathcal{P}(X)$ is the set of probability measures defined on $(X, \mathcal{B}(X))$, and $\mathcal{P}(X|Y)$ is the set of stochastic kernels on $X$ given $Y$ where $Y$ denotes a Borel space. When referring to the space of measures $\mathcal{M}(X)$, it is supposed that this space is endowed with the weak topology. Suppose that $Y = W \times Z$ where $W$ and $Z$ are Borel spaces. The marginal of a measure $\eta \in \mathcal{M}(Y)$ with respect to the first space $W$ will be denoted by $\widehat{\eta}$, that is, $\widehat{\eta}(\Gamma_W) = \eta(\Gamma_W \times Z)$ for any $\Gamma_W \in \mathcal{B}(W)$. For any point $x \in X$, $\delta_x$ denotes the Dirac measure defined by $\delta_x(\Gamma) = I_\Gamma(x)$ for any $\Gamma \in \mathcal{B}(X)$. The set of bounded real-valued measurable functions defined on the Borel space $X$ is denoted by $\mathbb{B}(X)$ and $\mathbb{C}(X)$ (respectively, $\mathbb{L}(X)$ and $\mathbb{U}(X)$) is the set of bounded real-valued continuous (respectively, lower semi-continuous and upper semi-continuous) functions defined on $X$. Finally, the infimum over an empty set is understood to be equal to $+\infty$, and we set $e^{-\infty} = 0$.

## 2.1 Parameters of the model

We will deal with a control model defined through the following elements:

- $\mathbf{X}$ is the state space, assumed to be an open subset of $\mathbb{R}^d$ ($d \in \mathbb{N}^*$) and $\partial\mathbf{X}$ denotes the boundary of $\mathbf{X}$.
- $\phi(x, t) : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ is the flow satisfying $\phi(x, t + s) = \phi(\phi(x, s), t)$ for all $x \in \mathbb{R}^d$ and $(t, s) \in \mathbb{R}^2$.
- $\boldsymbol{\Xi} = \{x \in \partial\mathbf{X} : x = \phi(y, t)$ for some $y \in \mathbf{X}$ and $t \in \mathbb{R}_+^*\}$ is the so called active boundary. Below, with some abuse of notation, $\overline{\mathbf{X}}$ denotes $\mathbf{X} \cup \boldsymbol{\Xi}$. For $x \in \overline{\mathbf{X}}$ we use the notation $t^*(x) = \inf\{t \in \mathbb{R}_+ : \ \phi(x, t) \in \boldsymbol{\Xi}\}$. Actually, the flow $\phi$ outside the space $\overline{\mathbf{X}}$ plays no role and can be defined arbitrarily.
- $\mathbf{A}$ is the action space, assumed to be a Borel space. $\mathbf{A}^i \in \mathcal{B}(\mathbf{A})$ (respectively $\mathbf{A}^g \in \mathcal{B}(\mathbf{A})$) is the set of impulsive (respectively gradual) actions satisfying $\mathbf{A} = \mathbf{A}^i \cup \mathbf{A}^g$ with $\mathbf{A}^i \cap \mathbf{A}^g = \emptyset$. Impulsive actions will be also called interventions.
- The set of feasible actions in state $x \in \overline{\mathbf{X}}$ is $\mathbf{A}(x) \in \mathcal{B}(\mathbf{A})$. We assume that $\mathbf{A}(x) \cap \mathbf{A}^i \neq \emptyset$ for all $x \in \boldsymbol{\Xi}$ and $\mathbf{A}(x) \cap \mathbf{A}^g \neq \emptyset$ for all $x \in \overline{\mathbf{X}}$. We introduce the set $\mathbf{K} = \mathbf{K}^i \cup \mathbf{K}^g$ with
$$\mathbf{K}^i = \{(x, a) \in \overline{\mathbf{X}} \times \mathbf{A} : a \in \mathbf{A}(x) \cap \mathbf{A}^i\} \in \mathcal{B}(\overline{\mathbf{X}}^i \times \mathbf{A}^i),$$
where $\overline{\mathbf{X}}^i = \{x \in \overline{\mathbf{X}} : \mathbf{A}(x) \cap \mathbf{A}^i \neq \emptyset\} \in \mathcal{B}(\overline{\mathbf{X}})$,
$$\mathbf{K}^g = \{(x, a) \in \overline{\mathbf{X}} \times \mathbf{A} : a \in \mathbf{A}(x) \cap \mathbf{A}^g\} \in \mathcal{B}(\overline{\mathbf{X}} \times \mathbf{A}^g),$$
It is assumed that $\mathbf{K}^i$ (respectively, $\mathbf{K}^g$) contains the graph of a measurable function from $\overline{\mathbf{X}}^i$ to $\mathbf{A}^i$ (respectively, from $\overline{\mathbf{X}}$ to $\mathbf{A}^g$). Below, $\mathbf{A}^i(x) = \mathbf{A}(x) \cap \mathbf{A}^i$ and $\mathbf{A}^g(x) = \mathbf{A}(x) \cap \mathbf{A}^g$ are admissible impulsive and gradual actions, correspondingly, in state $x \in \overline{\mathbf{X}}$.
- The controlled jump intensity $\lambda$ which is a $\mathbb{R}_+$-valued measurable function defined on $\mathbf{K}^g$.
- The stochastic kernel $Q$ on $\mathbf{X}$ given $\mathbf{K}$ satisfying $Q(\mathbf{X} \setminus \{x\}|x, a) = 1$ for any $(x, a) \in \mathbf{K}^g$. It describes the state of the process after any jump.

It should be noticed that in the framework of continuous-time MDPs, the signed kernel on $\mathbf{X}$ given $\mathbf{K}^g$, defined by
$$q(dy|x, a) = \lambda(x, a)\big[Q(dy|x, a) - \delta_x(dy)\big] \tag{1}$$
is the (controlled) infinitesimal generator of the jump process. For any $(z, b) \in \mathbf{K}^i$, $Q(\cdot|z, b)$ is the distribution of the state immediately after a jump either from the boundary or because of an

intervention. In other words, it describes the result of an impulsive action $b \in \mathbf{A}(z)$, when the state changes instantly. We call such jumps as 'forced' jumps, while the jumps governed by the generator $q$ are called 'natural' jumps.

Related to the parameters defining the process, one needs to introduce the set $\mathbb{A}(\overline{\mathbf{X}})$ of bounded measurable functions which are absolutely continuous with respect to the flow $\phi$, that is, the set of functions $g \in \mathbb{B}(\overline{\mathbf{X}})$ such that, for any $x \in \overline{\mathbf{X}}$, the function $g(\phi(x,\cdot))$ is absolutely continuous on $[0, t^*(x)] \cap \mathbb{R}_+$. If $g \in \mathbb{B}(\mathbf{X})$ is such that for any $x \in \mathbf{X}$, $g(\phi(x,\cdot))$ is absolutely continuous on $[0, t^*(x)[$ and $\lim_{t \to t^*(x)} g(\phi(x,t))$ exists whenever $t^*(x) < \infty$ then it can be easily seen that the domain of the mapping $g$ can be extended to $\overline{\mathbf{X}}$ by setting $g(z) = \lim_{t \to t^*(x)} g(\phi(x,t))$ where $z = \phi(x, t^*(x)) \in \mathbf{\Xi}$. By doing so, we can consider that $g \in \mathbb{A}(\overline{\mathbf{X}})$. Let $g \in \mathbb{A}(\overline{\mathbf{X}})$. From Lemma 2.2 in [3], there exists a real-valued measurable function $\mathcal{X}g$ defined on $\overline{\mathbf{X}}$ satisfying

$$g(\phi(x,t)) = g(x) + \int_{[0,t]} \mathcal{X}g(\phi(x,s))ds, \tag{2}$$

for any $x \in \overline{\mathbf{X}}$, $t \in [0, t^*(x)] \cap \mathbb{R}_+$. Observe that for any function $g \in \mathbb{A}(\overline{\mathbf{X}})$, the function $\mathcal{X}g$ satisfying (2) is not necessarily unique.

Suppose $s^*(\cdot)$ is such a measurable function on $\overline{\mathbf{X}}$ with values in $\overline{\mathbb{R}}_+$ that $s^*(x) \leq t^*(x)$. We denote by $\mathbb{A}_{s^*}(\overline{\mathbf{X}})$ the set of functions $g \in \mathbb{B}(\overline{\mathbf{X}})$ for which there exists a measurable function $\mathcal{X}g$ on $\overline{\mathbf{X}}$ satisfying equality (2) for all $x \in \overline{\mathbf{X}}$ and $t \in [0, s^*(x)] \cap \mathbb{R}_+$.

We consider in this paper that the control acts only on the intensity of the jump and on the transition kernel. The main difficulty in considering the control acting also on the flow comes from the fact that in such a situation, the time $t_*(x)$ which the flow takes to hit the boundary starting from $x$ and the first order differential operator $\mathcal{X}$ associated to the flow would depend on the control. This makes the problem much more complicated to solve and a several technical difficulties arise.

## 2.2 Construction of the process

Let $\mathbf{X}_\infty = \mathbf{X} \cup \{x_\infty\}$, where $x_\infty$ is an isolated artificial point corresponding to the case when no jumps occur in the future. For notational convenience, we put $\mathbf{A}^i_\Delta = \mathbf{A}^i \cup \{\Delta\}$ and $\mathbf{A}^i_\Delta(x) = \mathbf{A}^i(x) \cup \{\Delta\}$, where the isolated point $\Delta$ is a fictitious impulsive action meaning no intervention and $\mathbf{A}^i_\Delta(x_\infty) = \mathbf{A}^g_\Delta(x_\infty) = \{\Delta\}$. We extend the definition of $Q$ by setting $Q(\Gamma|x,\Delta) = \delta_x(\Gamma)$ for all $x \in \mathbf{X}_\infty$, $\Gamma \in \mathcal{B}(\mathbf{X}_\infty)$.

Let us introduce

$$\Omega_n = \mathbf{X} \times \mathbf{A}^i_\Delta \times \mathbf{X} \times (\mathbb{R}^*_+ \times \overline{\mathbf{X}} \times \mathbf{A}^i_\Delta \times \mathbf{X})^n \times (\{\infty\} \times \{x_\infty\} \times \{\Delta\} \times \{x_\infty\})^\infty.$$

The canonical space denoted by $\Omega$ is defined as

$$\Omega = \left[\bigcup_{n=1}^{\infty} \Omega_n\right] \bigcup \left[\mathbf{X} \times \mathbf{A}^i_\Delta \times \mathbf{X} \times (\mathbb{R}^*_+ \times \overline{\mathbf{X}} \times \mathbf{A}^i_\Delta \times \mathbf{X})^\infty\right]$$

and is endowed with its Borel $\sigma$-algebra denoted by $\mathcal{F}$. For notational convenience, $\omega \in \Omega$ will be represented as

$$\omega = (x_0, a^i_0, \overline{x}_0, \theta_1, x_1, a^i_1, \overline{x}_1, \theta_2, x_2, a^i_2, \overline{x}_2, \ldots).$$

Here, $x_0 \in \mathbf{X}$ is the initial state of the controlled process $X(\omega,t)$ with values in $\mathbf{X}$, defined below and $\overline{x}_0$ is the result of the initial intervention $a^i_0$. For $n \geq 1$, the component $\theta_n$ describes the interval of time up to the next jump. Two cases have to be considered. If the jump is natural, then $x_n$ gives the new state. One allows a possible intervention after a natural jump leading to a forced jump immediately after this natural jump: $a_n^i \in \mathbf{A}_\Delta^i$ will denote the impulsive action chosen by the decision maker and $\overline{x}_n$ will be the corresponding forced jump. If the jump is forced, then $x_n = \phi(\overline{x}_{n-1}, \theta_n) \in \overline{\mathbf{X}}$ will denote the value of the process just before the time of intervention, $a_n^i \in \mathbf{A}^i$ will correspond to the impulsive action, and $\overline{x}_n$ will be the associated forced jump. Observe that $\phi(\overline{x}_{n-1}, \theta_n) = x_n = \overline{x}_n$ may happen if the impulsive action $a_n^i \in \mathbf{A}^i$ did not change the state $x_n$ at the forced jump moment. This description is illustrated on Figure 1.

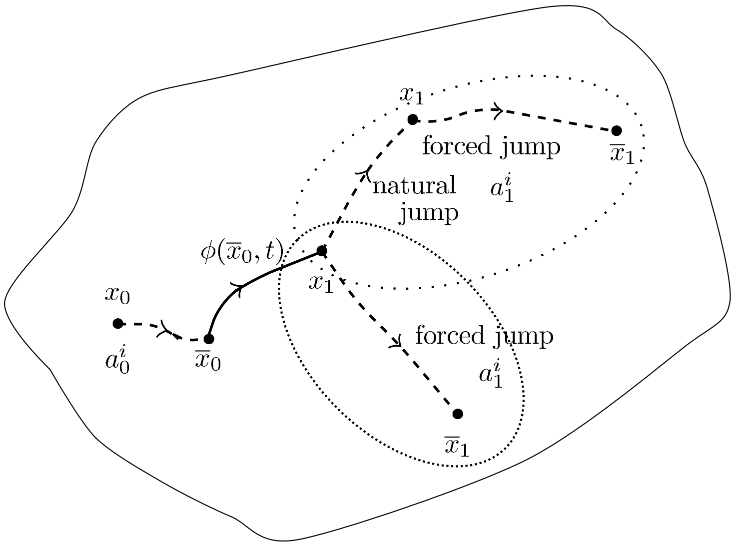


Figure 1: Dynamic of the controlled process. The two ovals describe the two different types of jumps from the state $x_1$.

To describe the situation where no jump occurs after the $n$-th jump, that is $\theta_n < \infty$ and $\theta_{n+1} = \infty$ (or just $\theta_1 = \infty$ for $n = 0$), i.e., the trajectory has only $n$ jumps, we set $\theta_m = \infty$, $x_m = \overline{x}_m = x_\infty$ (artificial point), $a_m^i = \Delta$ (artificial intervention) for all $m \geq n+1$. Between the jumps, the state of the process $X$ moves according to the flow $\phi$. The path up to the $n$-th jump is denoted by

$$h_n = (x_0, a_0^i, \overline{x}_0, \theta_1, x_1, a_1^i, \overline{x}_1, \theta_2, x_2, a_2^i, \overline{x}_2, \ldots, \theta_n, x_n, a_n^i, \overline{x}_n),$$

and the collection of all such paths is denoted by $\mathbf{H}_n$ ($n \in \mathbb{N}$). Additionally, $h_{0-} = x_0$ and $\mathbf{H}_{0-} = \mathbf{X}$. For $n \in \mathbb{N}$, introduce the mappings $X_n : \Omega \to \mathbf{X}_\infty \cup \mathbf{\Xi}$ and $\overline{X}_n : \Omega \to \mathbf{X}_\infty$ by $X_n(\omega) = x_n$ and $\overline{X}_n(\omega) = \overline{x}_n$. Similarly, for $n \geq 1$, the mapping $\Theta_n : \Omega \to \overline{\mathbb{R}}_+^*$ is defined by $\Theta_n(\omega) = \theta_n$ and $A_n^i(\omega) = a_n^i$. The sequence $(T_n)_{n \in \mathbb{N}^*}$ of $\overline{\mathbb{R}}_+^*$-valued mappings is defined on $\Omega$ by $T_n(\omega) = \sum_{i=1}^n \Theta_i(\omega) = \sum_{i=1}^n \theta_i$ and $T_\infty(\omega) = \lim_{n\to\infty} T_n(\omega)$. We denote by

$$H_n = (X_0, A_0^i, \overline{X}_0, \Theta_1, X_1, A_1^i, \overline{X}_1, \Theta_2, X_2, A_2^i, \overline{X}_2, \ldots, \Theta_n, X_n, A_n^i, \overline{X}_n)$$

the $n$-term random history taking values in $\mathbf{H}_n$ for $n \in \mathbb{N}$ and $H_{0-} = (X_0)$. The random elements are denoted with capital letters, the lowercase is for their realized values, that is, e.g., $t_n = \sum_{i=1}^n \theta_i$ is the realized jump moment and $\theta_i$ are the realized sojourn times.

The random measure $\mu$ associated with $(\Theta_n, X_n, A_n^i, \overline{X}_n)_{n\in\mathbb{N}^*}$ is a measure defined on $\mathbb{R}_+^* \times \overline{\mathbf{X}} \times \mathbf{A}_\Delta^i \times \mathbf{X}$ by

$$\mu(\omega; dt, dx, da, d\overline{x}) = \sum_{n\geq 1} I_{\{T_n(\omega)<\infty\}} \delta_{(T_n(\omega), X_n(\omega), A_n^i(\omega), \overline{X}_n(\omega))}(dt, dx, da, d\overline{x}).$$

For notational convenience the dependence on $\omega$ will be suppressed and, instead of $\mu(\omega; dt, dx, da, d\overline{x})$, it will be written $\mu(dt, dx, da, d\overline{x})$. For $t \in \mathbb{R}_+$, define

$$\mathcal{F}_t = \sigma\{H_0\} \vee \sigma\{\mu(]0, s] \times B) : \; s \leq t, B \in \mathcal{B}(\overline{\mathbf{X}} \times \mathbf{A}_\Delta^i \times \mathbf{X})\}.$$

Finally, we define the controlled process $\{X(t)\}_{t\in\mathbb{R}_+}$:

$$X(t,\omega) = \begin{cases} \phi(\overline{X}_n, t - T_n) & \text{if } T_n \leq t < T_{n+1} \text{ for } n \in \mathbb{N}; \\ x_\infty, & \text{if } T_\infty \leq t, \end{cases}$$

and $X(0-, \omega) = X_0 = x_0$. Below, we assume for simplicity that $x_0$ is fixed, but one can easily generalize the theory to an arbitrary initial distribution of the $X_0$ component.

### 2.3 Admissible strategies, probability measure and compensator

For $x \in \mathbf{X}$ with $t^*(x) < \infty$, we define

$$\mathbb{T}_x^i = \{t \in [0, t^*(x)] : \quad \phi(x,t) \in \overline{\mathbf{X}}^i\};$$

for $x \in \mathbf{X}$ with $t^*(x) = \infty$,

$$\mathbb{T}_x^i = \{t \in \mathbb{R}_+ : \quad \phi(x,t) \in \overline{\mathbf{X}}^i\} \cup \{\infty\},$$

and $\mathbb{T}_{x_\infty}^i = \{\infty\}$. $\mathbb{T}_x^i$ is the set of time moments, after a jump of the process into the state $x$, where the decision maker may apply an impulsive action. In the context of $t^*(x) = \infty$, observe that $\infty$ belongs to $\mathbb{T}_x^i$, corresponding to the situation where the decision maker chooses not to apply an impulsive action.

**Remark 2.1** *$\mathbb{T}_x^i$ is a subset of $\overline{\mathbb{R}}_+$. We consider $\overline{\mathbb{R}}_+$ as a compact metrizable space (adding $\infty$ as the one-point compactification, see Theorem 2.72 in [1]). In this context, we assume that the set $\mathbb{Y} = \{(x,t) : x \in \overline{\mathbf{X}}, t \in \mathbb{T}_x^i\}$ is closed in $\overline{\mathbf{X}} \times \overline{\mathbb{R}}_+$. By the way, $\mathbb{T}_x^i$ is closed (hence compact) for any $x \in \overline{\mathbf{X}}$.*

An admissible control strategy is a sequence $u = (u_n)_{n\in\mathbb{N}}$ such that $u_0$ is a stochastic kernel on $\mathbf{A}_\Delta^i$ given $\mathbf{X}$ with $u_0(\mathbf{A}_\Delta^i(x)|x) = 1$ and, for any $n \in \mathbb{N}^*$, $u_n$ is given by

$$u_n = \left(\psi_n, \pi_n, \gamma_n^0, \gamma_n^1\right),$$

where

- $\psi_n$ is a stochastic kernel on $\overline{\mathbb{R}}_+^*$ given $\mathbf{H}_{n-1}$ satisfying $\psi_n(\mathbb{T}_{\overline{x}_{n-1}}^i|h_{n-1}) = 1$ for any

$$h_{n-1} = (x_0, a_0^i, \overline{x}_0, \theta_1, x_1, a_1^i, \overline{x}_1, \theta_2, x_2, a_2^i, \overline{x}_2, \ldots, \theta_{n-1}, x_{n-1}, a_{n-1}^i, \overline{x}_{n-1}) \in \mathbf{H}_{n-1};$$

- $\pi_n$ is a stochastic kernel on $\mathbf{A}^g$ given $\mathbf{H}_{n-1} \times \mathbb{R}_+^*$. For $h_{n-1} \in \mathbf{H}_{n-1}$ with $\overline{x}_{n-1} \neq x_\infty$, it satisfies $\pi_n(\mathbf{A}^g(\phi(\overline{x}_{n-1}, t))|h_{n-1}, t) = 1$ for any $t \in \mathbb{R}_+^* \cap ]0, t^*(\overline{x}_{n-1})]$. In the case $\overline{x}_{n-1} = x_\infty$, $\pi_n$ is arbitrarily fixed.

- $\gamma_n^0$ and $\gamma_n^1$ are stochastic kernels on $\mathbf{A}^i_\Delta$ given $\mathbf{H}_{n-1}\times\overline{\mathbf{X}}$ satisfying $\gamma_n^0(\cdot|h_{n-1},y),\ \gamma_n^1(\cdot|h_{n-1},y)\in\mathcal{P}(\mathbf{A}^i_\Delta(y))$ for any $h_{n-1}\in\mathbf{H}_{n-1}$, $y\in\overline{\mathbf{X}}$.

The meaning of the elements of a strategy $u$ can be described as follows. Given the initial state $x_0$, $u_0$ is the initial randomized intervention with its corresponding realization $A^i_0$. For any $n\in\mathbb{N}^*$, after the jump of the process into the state $\overline{x}_{n-1}\in\mathbf{X}$ at the time moment $t_{n-1}\in\mathbb{R}_+$, the decision maker has to choose the following components:
– $\psi_n$, the distribution of the time interval $\tau$ up to the next intervention, concentrated on the set $\mathbb{T}^i_{\overline{x}_{n-1}}$ where interventions are possible;
– $\pi_n$, the distribution of the gradual control influencing the jump intensity that generates a natural jump;
– the distribution of the intervention through two different stochastic kernels: $\gamma^0$ and $\gamma^1$ depending on the fact that a natural or a forced jump has occurred. $\gamma^0$ corresponds to a natural jump and $\gamma^1$ is related to a forced jump. These distributions may depend on the state of the process before the intervention.

**Remark 2.2** *We underline that, since $\theta_n>0$ for all $n\in\mathbb{N}^*$, a sequence of simultaneous impulsive actions is not allowed. To guarantee such assumption, one can impose different conditions.*

*(a) For any $(x,a)\in\mathbf{K}^i$, the $Q(\cdot|x,a)$ measure is concentrated on $\mathbf{X}\setminus\overline{\mathbf{X}}^i$.*

*(b) For any $(x,a)\in\mathbf{K}^i$, the measure $Q(\Gamma|x,a)=\delta_{l(x,a)}(\Gamma)$ is degenerate and the new state $l(x,a)\neq x$ is such that, if $y=l(l(x,a_1),a_2)$, then there is $a_3\in\mathbf{A}^i(x)$ such that $y=l(x,a_3)$ and the associated costs satisfy $C^i(x,a_1)+C^i(l(x,a_1),a_2)>C^i(x,a_3)$. Here $C^i(x,a)$ is the cost of the impulsive action $a$ in state $x$; more about that in Section 3. Similar condition was imposed in [6].*

*(c) One can say that a finite simultaneous sequence of impulsive actions is one intervention and $C^i(x,a)>\delta>0$. This approach, demonstrated in [8, 9], causes no principal difficulties, but leads to cumbersome notations.*

**Remark 2.3** *Note that, if $u_0(\{\Delta\}|x_0)=1$ and $\psi_n(\{t^*(\overline{x}_{n-1})\}|h_{n-1})=1$ for all $n=1,2,\ldots$ then the forced jumps occur only from the boundary and all the results in [4] are valid for such strategies.*

Suppose a strategy $u=(u_n)_{n\in\mathbb{N}}\in\mathcal{U}$ is fixed with $u_n=\big(\psi_n,\pi_n,\gamma_n^0,\gamma_n^1\big)$ for $n\in\mathbb{N}^*$ and let $n\in\mathbb{N}^*$ be fixed. For $h_{n-1}\in\mathbf{H}_{n-1}$ with $\overline{x}_{n-1}\neq x_\infty$, we introduce the intensity of the natural jumps

$$\lambda_n^u(h_{n-1},t) \;=\; \int_{\mathbf{A}^g}\lambda(\phi(\overline{x}_{n-1},t),a)\pi_n(da|h_{n-1},t),\quad t\in\mathbb{R}^*_+\cap]0,t^*(\overline{x}_{n-1})],$$

and the rate of the natural jumps

$$\Lambda_n^u(h_{n-1},t) \;=\; \int_{]0,t]}\lambda_n^u(h_{n-1},s)ds,\quad t\in\mathbb{R}^*_+\cap]0,t^*(\overline{x}_{n-1})].$$

In case $\overline{x}_{n-1}=x_\infty$, $\lambda_n^u(h_{n-1},t)=\Lambda_n^u(h_{n-1},t)=0$ for all $t\in\mathbb{R}^*_+$.

Consider $n\in\mathbb{N}^*$ and $h_{n-1}\in\mathbf{H}_{n-1}$ with $\overline{x}_{n-1}\neq x_\infty$ and $\Gamma\in\mathcal{B}(\mathbf{X})$. According to the definition of the transition kernel $q(\cdot|x,a)$ (see equation (1)), we define the distribution of the state $X_n$ immediately after a natural jump at the time moment $\sum_{k=1}^{n-1}\theta_k+t$ for $t<t^*(x_{n-1})$ by

$$\frac{\int_{\mathbf{A}^g}q(\Gamma\setminus\{\phi(\overline{x}_{n-1},t)\})|\phi(\overline{x}_{n-1},t),a^g)\pi_n(da^g|h_{n-1},t)}{\lambda_n^u(h_{n-1},t)}$$

$$= \frac{\int_{\mathbf{A}^g} Q(\Gamma|\phi(\overline{x}_{n-1},t),a^g)\lambda(\phi(\overline{x}_{n-1},t),a^g)\pi_n(da^g|h_{n-1},t)}{\lambda_n^u(h_{n-1},t)}.$$

After that, an intervention $A_n^i$ may be applied immediately leading to the following distribution of $(X_n, A_n^i, \overline{X}_n)$:

$$\begin{aligned}Q_n^{g,u}(\Gamma_{\mathbf{X}} \times \Gamma_{\mathbf{A}} \times \overline{\Gamma}_{\mathbf{X}}|h_{n-1},t) =& \frac{1}{\lambda_n^u(h_{n-1},t)} \int_{\mathbf{A}^g} \bigg[ \int_{\Gamma_{\mathbf{X}}} \bigg( \int_{\Gamma_{\mathbf{A}}} Q(\overline{\Gamma}_{\mathbf{X}}|y,a^i)\gamma_n^0(da^i|h_{n-1},y) \bigg) \\ & \times Q(dy|\phi(\overline{x}_{n-1},t),a^g) \bigg] \lambda(\phi(\overline{x}_{n-1},t),a^g)\pi_n(da^g|h_{n-1},t)\end{aligned}$$

for $\Gamma_{\mathbf{X}} \in \mathcal{B}(\overline{\mathbf{X}}), \overline{\Gamma}_{\mathbf{X}} \in \mathcal{B}(\mathbf{X})$, $\Gamma_{\mathbf{A}} \in \mathcal{B}(\mathbf{A}_\Delta^i)$. In case $\lambda_n^u(h_{n-1},t) = 0$, $Q_n^{g,u}$ is fixed arbitrarily. When $\overline{x}_{n-1} \neq x_\infty$ and $\tau \neq \infty$, one can introduce the distribution of $(X_n, A_n^i, \overline{X}_n)$ after a forced jump at time moment $\sum_{k=1}^{n-1} \theta_k + \tau$ with $\tau \in \mathbb{T}_{\overline{x}_{n-1}}^i$ by

$$\begin{aligned}Q_n^{i,u}(\Gamma_{\mathbf{X}} \times \Gamma_{\mathbf{A}} \times \overline{\Gamma}_{\mathbf{X}}|h_{n-1},\tau) =& \delta_{\phi(\overline{x}_{n-1},\tau-)}(\Gamma_{\mathbf{X}}) \int_{\Gamma_{\mathbf{A}}} Q(\overline{\Gamma}_{\mathbf{X}}|\phi(\overline{x}_{n-1},\tau-),a^i) \\ & \times \gamma_n^1(da^i|h_{n-1},\phi(\overline{x}_{n-1},\tau-)),\end{aligned}$$

for $\Gamma_{\mathbf{X}} \in \mathcal{B}(\overline{\mathbf{X}}), \overline{\Gamma}_{\mathbf{X}} \in \mathcal{B}(\mathbf{X})$ and $\Gamma_{\mathbf{A}} \in \mathcal{B}(\mathbf{A}_\Delta^i)$. In case $\overline{x}_{n-1} = x_\infty$ or $\tau = \infty$, or $\tau \notin \mathbb{T}_{\overline{x}_{n-1}}^i$, the distribution $Q_n^{i,u}$ is fixed arbitrarily.

Now, for any $n \in \mathbb{N}^*$, the stochastic kernel $G_n$ on $\overline{\mathbb{R}}_+^* \times (\mathbf{X}_\infty \cup \mathbf{\Xi}) \times \mathbf{A}_\Delta^i \times \mathbf{X}_\infty$ given $\mathbf{H}_{n-1}$, describing the joint distribution of $(\Theta_n, X_n, A_n^i, \overline{X}_n)$, is defined by

$$G_n(\{+\infty\} \times \{x_\infty\} \times \Delta \times \{x_\infty\}|h_{n-1}) = \delta_{\overline{x}_{n-1}}(\{x_\infty\}) + \delta_{\overline{x}_{n-1}}(\mathbf{X})e^{-\Lambda_n^u(h_{n-1},\infty)}\psi_n(\{+\infty\}|h_{n-1}) \tag{3}$$

and

$$\begin{aligned}G_n(\Gamma_\Theta \times \Gamma|h_{n-1}) =& \delta_{\overline{x}_{n-1}}(\mathbf{X}) \bigg[ \int_{\Gamma_\Theta} Q_n^{i,u}(\Gamma|h_{n-1},\tau)e^{-\Lambda_n^u(h_{n-1},\tau)}\psi_n(d\tau|h_{n-1}) \\ & + \int_{\Gamma_\Theta} \psi_n([t,\infty]|h_{n-1})Q_n^{g,u}(\Gamma|h_{n-1},t)\lambda_n^u(h_{n-1},t)e^{-\Lambda_n^u(h_{n-1},t)}dt \bigg],\end{aligned} \tag{4}$$

where $\Gamma \in \mathcal{B}(\overline{\mathbf{X}} \times \mathbf{A}_\Delta^i \times \mathbf{X})$, $\Gamma_\Theta \in \mathcal{B}(\mathbb{R}_+^*)$ and $h_{n-1} \in \mathbf{H}_{n-1}$. Note that the kernels $Q_n^{i,u}$ and $Q_n^{g,u}$ appear in the formula for $G_n$ only if $\tau \in \mathbb{T}_{\overline{x}_{n-1}}^i$ and $\lambda_n^u(h_{n-1},t) \neq 0$ correspondingly.

Consider a strategy $u$ and an initial state $x_0 \in \mathbf{X}$. From Remark 3.43 in [13], there exists a probability $\mathbb{P}_{x_0}^u$ on $(\Omega, \mathcal{F})$ such that the restriction of $\mathbb{P}_{x_0}^u$ to $(\Omega, \mathcal{F}_0)$ is given by

$$\mathbb{P}_{x_0}^u\big(\{x_0\} \times \Gamma_{\mathbf{A}} \times \overline{\Gamma}_{\mathbf{X}} \times (\overline{\mathbb{R}}_+^* \times \mathbf{X}_\infty \times \mathbf{A}_\Delta^i \times \mathbf{X}_\infty)^\infty\big) \ = \ \int_{\Gamma_{\mathbf{A}}} Q(\overline{\Gamma}_{\mathbf{X}}|x_0,a^i)u_0(da^i|x_0) \tag{5}$$

for any $\Gamma_{\mathbf{A}} \in \mathcal{B}(\mathbf{A}_\Delta^i)$, $\overline{\Gamma}_{\mathbf{X}} \in \mathcal{B}(\mathbf{X})$, and the positive random measure $\nu$ defined on $\mathbb{R}_+^* \times \overline{\mathbf{X}} \times \mathbf{A}_\Delta^i \times \mathbf{X}$ by

$$\nu(dt,dx,da,d\overline{x}) = \sum_{n \in \mathbb{N}^*} \frac{G_n(dt - T_{n-1},dx,da,d\overline{x}|H_{n-1})}{G_n([t - T_{n-1},+\infty] \times \mathbf{X}_\infty \times \mathbf{A}_\Delta^i \times \mathbf{X}_\infty|H_{n-1})} I_{\{T_{n-1}<t\leq T_n\}} \tag{6}$$

is the predictable projection of $\mu$ with respect to $\mathbb{P}_{x_0}^u$.

According to Remark 2.2, we consider only "admissible" strategies $u$ for which $\mathbb{P}_{x_0}^u(\Theta_n = 0) = 0$ for all $n \in \mathbb{N}^*$. Below, the class of admissible control strategies is denoted by $\mathcal{U}$.

**Remark 2.4** *Observe that $\mathcal{F}_{T_n}$ is the $\sigma$-algebra generated by the random variable $H_n$ for $n \in \mathbb{N}$. The conditional distribution of $(\Theta_n, X_n, A_n^i, \overline{X}_n)$ given $\mathcal{F}_{T_{n-1}}$ for $n \in \mathbb{N}^*$ under $\mathbb{P}_{x_0}^u$ is determined by $G_n(\cdot|H_{n-1})$ and the conditional survival function of $\Theta_n$ given $\mathcal{F}_{T_{n-1}}$ under $\mathbb{P}_{x_0}^u$ is given by $G_n([t,+\infty] \times (\mathbf{X}_\infty \cup \mathbf{\Xi}) \times \mathbf{A}_\Delta^i \times \mathbf{X}_\infty|H_{n-1})$.*

In Lemma 2.5 below we prove the decomposition of the predictable projection $\nu$ of the process in terms of two parts: one being related to the natural jumps governed by the jump intensity $\lambda$ and the other to the forced jumps.

**Lemma 2.5** *For $u \in \mathcal{U}$, the predictable projection of the random measure $\mu$ is given by $\nu = \nu_0 + \nu_1$, where, for $\Gamma \in \mathcal{B}(\mathbb{R}_+^*)$, $\Gamma_{\mathbf{X}} \in \mathcal{B}(\overline{\mathbf{X}}), \overline{\Gamma}_{\mathbf{X}} \in \mathcal{B}(\mathbf{X})$, $\Gamma_{\mathbf{A}} \in \mathcal{B}(\mathbf{A}_\Delta^i)$,*

$$
\begin{aligned}
&\nu_0(\Gamma \times \Gamma_{\mathbf{X}} \times \Gamma_{\mathbf{A}} \times \overline{\Gamma}_{\mathbf{X}}) \\
=\ & \sum_{n\in\mathbb{N}^*} \int_\Gamma I_{\{T_{n-1}<s\leq T_n\}} \int_{\mathbf{A}^g} \left[ \int_{\Gamma_{\mathbf{X}}} \left[ \int_{\Gamma_{\mathbf{A}}} Q(\overline{\Gamma}_{\mathbf{X}}|y,a^i)\gamma_n^0(da^i|H_{n-1},y) \right] \right. \\
& \left. \times Q(dy|X(s),a^g) \right] \lambda(X(s),a^g)\pi_n(da^g|H_{n-1},s-T_{n-1})ds \\
=\ & \int_\Gamma \int_{\mathbf{A}^g} \left[ \int_{\Gamma_{\mathbf{X}}} \left[ \int_{\Gamma_{\mathbf{A}}} Q(\overline{\Gamma}_{\mathbf{X}}|y,a^i)\gamma^0(da^i|y,s) \right] Q(dy|X(s),a^g) \right] \lambda(X(s),a^g)\pi(da^g|s)ds, \\
&\nu_1(\Gamma \times \Gamma_{\mathbf{X}} \times \Gamma_{\mathbf{A}} \times \overline{\Gamma}_{\mathbf{X}}) \\
=\ & \sum_{n\in\mathbb{N}^*} \int_\Gamma I_{\{T_{n-1}<s\leq T_n\}} \delta_{X(s-)}(\Gamma_{\mathbf{X}}) \int_{\Gamma_{\mathbf{A}}} Q(\overline{\Gamma}_{\mathbf{X}}|X(s-),a^i)\gamma_n^1(da^i|H_{n-1},X(s-)) \frac{\psi_n(ds-T_{n-1}|H_{n-1})}{\psi_n([s-T_{n-1},\infty]|H_{n-1})}.
\end{aligned}
$$

*Here and below,*

$$
\pi(da^g|s) = \sum_{n\in\mathbb{N}^*} I_{\{T_{n-1}<s\leq T_n\}}\pi_n(da^g|H_{n-1},s-T_{n-1});
$$

$$
\gamma^0(da^i|y,s) = \sum_{n\in\mathbb{N}^*} I_{\{T_{n-1}<s\leq T_n\}}\gamma_n^0(da^i|H_{n-1},y).
$$

**Proof:** First observe that by using the integration by parts formula, we obtain that

$$
\begin{aligned}
& G_n([t,+\infty] \times (\mathbf{X}_\infty \cup \mathbf{\Xi}) \times \mathbf{A}_\Delta^i \times \mathbf{X}_\infty|h_{n-1}) \\
=\ & \delta_{\overline{x}n-1}(\{x_\infty\}) + \delta_{\overline{x}_{n-1}}(\mathbf{X})e^{-\Lambda_n^u(h_{n-1},t)}\psi_n([t,\infty]|h_{n-1}).
\end{aligned}
$$

Now, recalling the definition of $\nu$ (see equation (6)) in terms of $G$ (see equation (4)), the straightforward calculation gives the result. □

Note, $\nu_0(\mathbb{R}_+^* \times \mathbf{\Xi} \times \mathbf{A}_\Delta^i \times \mathbf{X}) = 0$.

# 3 Optimization problem and assumptions

## 3.1 Formulation of the optimization control problem

The objective of this subsection is to introduce the infinite-horizon performance criterion we are concerned with, as well as several different classes of admissible strategies.

The cost rate $C^g$ associated with a gradual action is a real-valued function defined on $\mathbf{K}^g$ and the cost $C^i$ associated with an impulsive action is a real-valued function defined on $\mathbf{K}^i$, and

$C^i(x,\Delta)=0$ for all $x\in\mathbf{X}$: no intervention - no cost. The associated infinite-horizon discounted criterion corresponding to an admissible control strategy $u\in\mathcal{U}$ is defined by

$$\begin{aligned}
\mathcal{V}(u,x_0) &= \overline{\lim_{N\to\infty}} \Bigg\{ \mathbb{E}^u_{x_0}\Bigg[\int_{[0,T_N]\cap\mathbb{R}_+} e^{-\alpha s}\int_{\mathbf{A}^g(X(s))} C^g(X(s),a^g)\pi(da^g|s)ds\Bigg] \\
&\quad + \mathbb{E}^u_{x_0}\Bigg[\sum_{n=0}^{N} e^{-\alpha T_n} C^i(X_n,A^i_n)\Bigg]\Bigg\} \\
&= \overline{\lim_{N\to\infty}} \Bigg\{ \mathbb{E}^u_{x_0}\Bigg[\int_{[0,T_N]\cap\mathbb{R}_+} e^{-\alpha s}\int_{\mathbf{A}^g(X(s))} C^g(X(s),a^g)\pi(da^g|s)ds\Bigg] \\
&\quad + \int_{\mathbf{A}^i_\Delta(x_0)} C^i(x_0,a^i)u_0(da^i|x_0) \\
&\quad + \mathbb{E}^u_{x_0}\Bigg[\int_{([0,T_N]\cap\mathbb{R}_+)\times\overline{\mathbf{X}}\times\mathbf{A}^i_\Delta} e^{-\alpha s}C^i(x,a^i)\mu(ds,dx,da^i,\mathbf{X})\Bigg]\Bigg\}.
\end{aligned} \tag{7}$$

Here $\alpha>0$ is the discount factor. Note that, for any control strategy $u\in\mathcal{U}$, the function $\mathcal{V}(u,\cdot)$ is measurable.

**Definition 3.1** *The optimization problem consists in minimizing the performance criterion $\mathcal{V}(u,x_0)$ within the class of admissible strategies $u\in\mathcal{U}$, where $x_0$ is the initial state.*

We introduce now several different classes of admissible strategies that will be considered along the paper. A control strategy $u\in\mathcal{U}$ is called

- *stationary non-randomized*, if $u_0(\cdot|y)=\delta_{\varphi^i(y)}(\cdot)$; $\psi_n(\cdot|h_{n-1})=\delta_{\psi^s(\overline{x}_{n-1})}(\cdot)$, $\pi_n(\cdot|h_{n-1},t)=\delta_{\varphi^g(\phi(\overline{x}_{n-1},t))}(\cdot)$ and $\gamma^0_n(\cdot|h_{n-1},y)=\gamma^1_n(\cdot|h_{n-1},y)=\delta_{\varphi^i(y)}(\cdot)$ for $y\in\overline{\mathbf{X}}$, where $\psi^s$ is a measurable function from $\mathbf{X}_\infty$ to $\overline{\mathbb{R}}^*_+$ satisfying $\psi^s(x)\in\mathbb{T}^i_x$ for all $x\in\mathbf{X}_\infty$; $\varphi^g:\overline{\mathbf{X}}\to\mathbf{A}^g$ is a measurable mapping satisfying $\varphi^g(x)\in\mathbf{A}^g(x)$ for any $x\in\overline{\mathbf{X}}$; $\varphi^i:\overline{\mathbf{X}}\to\mathbf{A}^i_\Delta$ is a measurable mapping satisfying $\varphi^i(y)\in\mathbf{A}^i_\Delta(y)$ for any $y\in\overline{\mathbf{X}}$.
- *uniformly or persistently* optimal, if $u$ satisfies $\mathcal{V}(u,x_0)=\inf_{v\in\mathcal{U}}\mathcal{V}(v,x_0)$ simultaneously for all $x_0\in\mathbf{X}$ and hence for any initial distribution.

## 3.2 Assumptions

In this subsection we present a list of assumptions that we will consider along the paper. Assumptions A will be mainly used to show that the process is non-explosive and to provide an upper bound for the sum of the expected values of $e^{-\alpha T_n}$. Assumptions B and C will be mainly required to obtain the existence of an optimal selector (as defined in [12, Appendix D]) for the optimization problem.

**Assumption A.** There are constants $K_1,K_2,K_3,K_4>0$ and $\varepsilon>0$ such that $K_4<K_2$ and

*(A1) For any $(x,a)\in\mathbf{K}^g$, $\lambda(x,a)\le K_1$.*

*(A2) For all $x\in\mathbf{\Xi}$, $a\in\mathbf{A}^i(x)$, $Q(A_\varepsilon|x,a)=1$, where*

$$A_\varepsilon=\{z\in\mathbf{X}:t^*(z)>\varepsilon\}.$$

*(A3)* $C^i(x,a) \in (K_2, K_3)$ *for all* $x \in \mathbf{X}$, $a \in \mathbf{A}^i(x)$, *and* $|C^g(x,a)| \le K_3$ *for all* $(x,a) \in \mathbf{K}^g$; $|C^i(x,a)| \le K_4$ *for all* $x \in \boldsymbol{\Xi}$, $a \in \mathbf{A}^i(x)$.

**Assumption B.**

*(B1)* *The sets* $\mathbf{A}^g(y)$ *and* $\mathbf{A}^i(y)$ *are compact for every* $y \in \overline{\mathbf{X}}$.

*(B2)* *The kernel* $Q$ *is weakly continuous.*

*(B3)* *The function* $\lambda$ *is continuous on* $\mathbf{K}^g$.

*(B4)* *The flow* $\phi$ *is continuous on* $\mathbb{R}^d \times \mathbb{R}_+$.

*(B5)* *The function* $t^*$ *is continuous on* $\overline{\mathbf{X}}$.

*(B6)* *The set* $\mathbb{Y} = \{(x,t) : x \in \overline{\mathbf{X}}, t \in \mathbb{T}^i_x\}$ *is closed in* $\overline{\mathbf{X}} \times \overline{\mathbb{R}}_+$, $\overline{\mathbb{R}}_+$ *is compact. (See Remark 2.1.)*

**Assumption C.**

*(C1)* *The multifunction* $\Psi^g$ *from* $\overline{\mathbf{X}}$ *to* $\mathbf{A}$ *defined by* $\Psi^g(x) = \mathbf{A}^g(x)$ *is upper semi-continuous. The multifunction* $\Psi^i$ *from* $\overline{\mathbf{X}}^i$ *to* $\mathbf{A}$ *defined by* $\Psi^i(z) = \mathbf{A}^i(z)$ *is upper semi-continuous.*

*(C2)* *The multifunction* $x \to \{t \in \overline{\mathbb{R}}_+ : (x,t) \in \mathbb{Y}\}$ *is upper semi-continuous.*

*(C3)* *The cost function* $C^g$ *(respectively,* $C^i$*) is bounded and lower semi-continuous on* $\mathbf{K}^g$ *(respectively,* $\mathbf{K}^i$*).*

# 4 Preliminary results

In this section we establish some preliminary results that will be needed throughout the paper. We start with Lemma 4.1 by showing that the process is non-explosive and provide an upper bound for the sum of the expected values of $e^{-\alpha T_n}$. The key result will be the discounted version of the so-called Dynkin formula associated with the controlled process, proved in Lemma 4.3.

**Lemma 4.1** *Suppose Assumptions (A1), (A2), and (A3) are satisfied. Then the following statements hold for an arbitrarily fixed* $x_0$.

*(a) There is a strategy* $u \in \mathcal{U}$ *such that the performance criterion is finite:*

$$\mathcal{V}(u, x_0) \le \frac{K_3}{\alpha} + K_3 \frac{2e^{\alpha K_1 \varepsilon}}{1 - e^{-\alpha\varepsilon}} < \infty.$$

*(b) For any control strategy* $u \in \mathcal{U}$ *with a finite value of the performance criterion, there exists* $M < \infty$ *such that*

$$\mathbb{E}^u_{x_0}\Big[\sum_{n \in \mathbb{N}} e^{-\alpha T_n}\Big] \le M \text{ and } \mathbb{P}^u_{x_0}(T_\infty < +\infty) = 0.$$

**Proof:** (a) Consider an arbitrary strategy $u \in \mathcal{U}$ with $u_0(\{\Delta\}|x_0) = 1$ and $\psi_n(\{t^*(\overline{x}_{n-1})\}|h_{n-1}) = 1$ for all $n \in \mathbb{N}^*$. According to Lemma 4.1 in [4] (see also Remark 2.3), $\mathbb{E}^u_{x_0}\Big[\sum_{n \in \mathbb{N}} e^{-\alpha T_n}\Big] \le \dfrac{2e^{\alpha K_1 \varepsilon}}{1 - e^{-\alpha\varepsilon}} < \infty$ and, due to (7), the statement (a) follows.

(b) Suppose $u \in \mathcal{U}$ is a strategy satisfying $\mathcal{V}(u, x_0) \le V < \infty$. The sum $\mathbb{E}^u_{x_0}\Big[\sum_{n \in \mathbb{N}^*} e^{-\alpha T_n}\Big]$ is decomposed into five terms corresponding to the following types of jumps of the process $\{X(t)\}_{t \in \mathbb{R}_+}$:

1) natural jumps,
2) forced jumps from the boundary $\boldsymbol{\Xi}$ following natural jumps,

3) forced jumps from the boundary $\mathbf{\Xi}$ following forced jumps from the boundary,
4) forced jumps from the boundary $\mathbf{\Xi}$ following forced jumps from $\mathbf{X}$,
5) forced jumps from $\mathbf{X}$.

Observe that $T_n$ corresponds either to a natural jump if and only if $X_n \neq \phi(\overline{X}_{n-1}, \Theta_n)$ or to a forced jump at the boundary if and only if $X_n \in \mathbf{\Xi}$ or to a forced jump from $\mathbf{X}$ if and only if $X_n = \phi(\overline{X}_{n-1}, \Theta_n-) \in \mathbf{X}$. Consequently, $\mathbb{E}^u_{x_0}\Big[\sum_{n\in\mathbb{N}^*} e^{-\alpha T_n}\Big] \leq 1 + \sum_{k=1}^{5} \mathcal{S}_k$ where

$$\begin{aligned}
\mathcal{S}_1 &= \mathbb{E}^u_{x_0}\Big[\sum_{n\geq 2} e^{-\alpha T_n} I_{\{X_n \neq \phi(\overline{X}_{n-1},\Theta_n-)\}}\Big],\\
\mathcal{S}_2 &= \mathbb{E}^u_{x_0}\Big[\sum_{n\geq 2} e^{-\alpha T_n} I_{\{X_n \in \mathbf{\Xi}\}} I_{\{X_{n-1} \neq \phi(\overline{X}_{n-2},\Theta_{n-1}-)\}}\Big],\\
\mathcal{S}_3 &= \mathbb{E}^u_{x_0}\Big[\sum_{n\geq 2} e^{-\alpha T_n} I_{\{X_n \in \mathbf{\Xi}\}} I_{\{X_{n-1} \in \mathbf{\Xi}\}}\Big],\\
\mathcal{S}_4 &= \mathbb{E}^u_{x_0}\Big[\sum_{n\geq 2} e^{-\alpha T_n} I_{\{X_n \in \mathbf{\Xi}\}} I_{\{X_{n-1} = \phi(\overline{X}_{n-2},\Theta_{n-1}-)\}} I_{\{X_{n-1}\in\mathbf{X}\}}\Big],\\
\mathcal{S}_5 &= \mathbb{E}^u_{x_0}\Big[\sum_{n\geq 2} e^{-\alpha T_n} I_{\{X_n = \phi(\overline{X}_{n-1},\Theta_n-)\}} I_{\{X_n\in\mathbf{X}\}}\Big],
\end{aligned}$$

For the first term, we have

$$\begin{aligned}
\mathcal{S}_1 &\leq \mathbb{E}^u_{x_0}\Big[\int_{\mathbb{R}^*_+\times\overline{\mathbf{X}}\times\mathbf{A}^i_\Delta\times\mathbf{X}} e^{-\alpha s} I_{\{X(s-)\neq x\}} \mu(ds,dx,da,d\overline{x})\Big]\\
&= \mathbb{E}^u_{x_0}\Big[\int_{\mathbb{R}^*_+\times\overline{\mathbf{X}}\times\mathbf{A}^i_\Delta\times\mathbf{X}} e^{-\alpha s} I_{\{X(s-)\neq x\}} \nu_0(ds,dx,da,d\overline{x})\Big]\\
&= \mathbb{E}^u_{x_0}\Big[\sum_{n\geq 1}\int_{\mathbb{R}^*_+\times\overline{\mathbf{X}}\times\mathbf{A}^i_\Delta\times\mathbf{X}} e^{-\alpha s} I_{\{T_{n-1}<s\leq T_n\}} I_{\{\phi(\overline{X}_{n-1},s-T_{n-1}-)\neq x\}} \nu_0(ds,dx,da,d\overline{x})\Big]\\
&= \mathbb{E}^u_{x_0}\Big[\sum_{n\geq 1}\int_{\mathbb{R}^*_+} e^{-\alpha s} I_{\{T_{n-1}<s\leq T_n\}} \int_{\mathbf{A}^g} \lambda(X(s),a^g)\pi_n(da^g|H_{n-1},s-T_{n-1})ds\Big]\\
&\leq K_1 \int_{\mathbb{R}^*_+} e^{-\alpha s} ds = \frac{K_1}{\alpha}.
\end{aligned}$$

Since the jumps of the second type occur later than the previous natural jumps, we easily obtain

$$\mathcal{S}_2 \leq \mathcal{S}_1 \leq \frac{K_1}{\alpha}.$$

For the third term, consider the process $\{Y(t)\}_{t\in\mathbb{R}_+}$ given by

$$Y(t) = \sum_{k=3}^{\infty} T_{k-1} I_{[T_{k-1},T_k[}(t) I_{\{X(T_{k-2})\in\mathbf{\Xi}\}} I_{\{X(T_{k-1})\in\mathbf{\Xi}\}}.$$

It is an $\{\mathcal{F}_t\}$-adapted right continuous process with left limits. Therefore, $\{Y(t)\}_{t\in\mathbb{R}_+}$ is an $\{\mathcal{F}_t\}$-progressively measurable process and so, $\mathcal{Y} = \{(t,\omega) : Y(t,\omega) \geq t\}$ is a progressively measurable set. Define $S_k = \inf\big\{t \in \mathbb{R}_+ : [0,t] \cap \mathcal{Y} \text{ contains at least k+1 points}\big\}$ for $k \in \mathbb{N}^*$. From Corollary

6.17 in [10], it is an increasing sequence of $\{\mathcal{F}_t\}$-stopping times. Clearly, $S_k$ covers all the jumps of the process $\{X(t)\}_{t\in\mathbb{R}_+}$ corresponding to a forced jump from the boundary that follows also a forced jump from the boundary. Therefore,

$$\mathcal{S}_3 \leq \mathbb{E}^u_{x_0}\Big[\sum_{k\geq 1} e^{-\alpha S_k}\Big].$$

Observe that on $\{S_k = T_n\} \cap \{S_k < \infty\}$ we have $S_{k-1} \leq T_{n-1}$ and $S_k = T_{n-1} + \theta_n = T_{n-1} + t^*(\overline{X}_{n-1}) \geq T_{n-1} + \varepsilon$ by Assumption (A2). As a consequence, $S_k \geq S_{k-1} + \epsilon$ on $\{S_k = T_n\} \cap \{S_k < \infty\}$ implying that $S_k \geq S_{k-1} + \varepsilon$ and so

$$\mathcal{S}_3 \leq \mathbb{E}^u_{x_0}\Big[\sum_{k\geq 1} e^{-\alpha S_k}\Big] \leq \frac{1}{1-e^{-\alpha\varepsilon}}.$$

Recalling Assumption (A3), the total cost associated with the jumps of type 4 is greater than $(K_2 - K_4)\mathcal{S}_4$ and the total cost associated with the jumps type 5 is greater than $K_2\mathcal{S}_5 \geq (K_2 - K_4)\mathcal{S}_5$. Therefore, by Assumption (A3) it follows

$$\mathbb{E}^u_{x_0}\bigg[\int_{]0,+\infty[\times\overline{\mathbf{X}}\times\mathbf{A}^i_\Delta} e^{-\alpha s} C^i(x,a^i)\mu(ds,dx,da^i,\mathbf{X})\bigg]$$
$$\geq -K_3\mathcal{S}_1 - K_3\mathcal{S}_2 - K_3\mathcal{S}_3 + (K_2 - K_4)(\mathcal{S}_4 + \mathcal{S}_5).$$

Moreover,

$$\mathbb{E}^u_{x_0}\bigg[\int_{]0,+\infty[} e^{-\alpha s}\int_{\mathbf{A}^g(X(s))} C^g(X(s),a^g)\pi(da^g|s)ds\bigg] \geq -\frac{K_3}{\alpha}$$

and

$$\int_{\mathbf{A}^i_\Delta(x_0)} C^i(x_0,a^i)u_0(da^i|x_0) \geq -K_4.$$

Therefore, if $\mathcal{V}(u,x_0) \leq V$ then according to equation (7)

$$S_4 + S_5 \leq \frac{V + K_4 + \frac{K_3}{\alpha} + \frac{2K_3K_1}{\alpha} + \frac{K_3}{1-e^{-\alpha\varepsilon}}}{K_2 - K_4} = M_1,$$

and the first statement of the lemma holds for

$$M = M_1 + \frac{2K_1}{\alpha} + \frac{1}{1-e^{-\alpha\varepsilon}} + 2.$$

The last summand 2 corresponds to $T_0 = 0$ and to the possible first jump from the boundary which has no type. Notice that, if $\mathbb{P}^u_{x_0}(T_\infty < +\infty) > 0$ then $\mathbb{E}^u_{x_0}\Big[\sum_{n\in\mathbb{N}} e^{-\alpha T_n}\Big] = \infty$ showing the second statement. □

**Remark 4.2** *Assumption (A3) implies that $\mathcal{V}(u,x_0) \geq -\frac{K_3}{\alpha} > -\infty$ for any control strategy $u \in \mathcal{U}$.*

The following lemma provides the discounted version of the so-called Dynkin formula associated with the controlled process $X(t)$.

**Lemma 4.3** *Suppose a strategy $u \in \mathcal{U}$ is fixed and $s^* : \overline{\mathbf{X}} \to \overline{\mathbb{R}}_+$ is a measurable function such that $s^*(x) \leq t^*(x)$ for all $x \in \overline{\mathbf{X}}$. Let $W \in \mathbb{A}_{s^*}(\overline{\mathbf{X}})$ for which $\mathcal{X}W \in \mathbb{B}(\overline{\mathbf{X}})$, and $\alpha > 0$ be a discount factor. Then, we have for any $N \in \mathbb{N}^*$*

$$
\begin{aligned}
&\mathbb{E}^u_{x_0}\left[e^{-\alpha T_N} W(X(T_N))\right] \\
&= \int_{\mathbf{A}^i_\Delta(x_0)} \int_{\mathbf{X}} W(y) Q(dy|x_0, a^i) u_0(da^i|x_0) + \mathbb{E}^u_{x_0}\Bigg[\sum_{n=1}^{N} \Bigg\{ \int_{]T_{n-1}, T_{n-1}+s^*(\overline{X}_{n-1}) \wedge \Theta_n] \cap \mathbb{R}_+} e^{-\alpha s} \\
&\quad \times \big[\mathcal{X}W(X(s)) - \alpha W(X(s))\big] ds + I_{\{\Theta_n > s^*(\overline{X}_{n-1})\}} \Big(e^{-\alpha T_n} W(X(T_n-)) \\
&\quad - e^{-\alpha(T_{n-1}+s^*(\overline{X}_{n-1}))} W(X(T_{n-1} + s^*(\overline{X}_{n-1})-))\Big) \Bigg\} \Bigg] \\
&\quad + \mathbb{E}^u_{x_0}\Bigg[\int_{[0,T_N] \cap \mathbb{R}_+} \int_{\mathbf{A}^g(X(s))} \Bigg[\int_{\mathbf{X}} \Bigg[\int_{\mathbf{A}^i_\Delta(y)} \int_{\mathbf{X}} e^{-\alpha s} \Big[W(z) - W(X(s))\Big] Q(dz|y, a^i) \gamma^0(da^i|y, s)\Bigg] \\
&\quad \times\ Q(dy|X(s), a^g)\Bigg] \lambda(X(s), a^g) \pi(da^g|s) ds\Bigg] \\
&\quad + \mathbb{E}^u_{x_0}\Bigg[\sum_{n=1}^{N} \int_{]T_{n-1}, T_n] \cap \mathbb{R}_+} \int_{\mathbf{A}^i(X(s-))} e^{-\alpha s} \int_{\mathbf{X}} \Big[W(z) - W(X(s-))\Big] Q(dz|X(s-), a^i) \\
&\quad \times\ \gamma^1_n(da^i|H_{n-1}, X(s-)) \frac{\psi_n(ds - T_{n-1}|H_{n-1})}{\psi_n([s - T_{n-1}, \infty]|H_{n-1})}\Bigg].
\end{aligned} \tag{8}
$$

**Proof:** The proof is similar to that of Lemma 4.2 in [4]. It is easy to show

$$
\begin{aligned}
\mathbb{E}^u_{x_0}\left[e^{-\alpha T_N} W(X(T_N))\right] =& \mathbb{E}^u_{x_0}\big[W(\overline{X}(0))\big] + \mathbb{E}^u_{x_0}\Bigg[\sum_{n=1}^{N} \Bigg\{ \int_{]T_{n-1}, T_{n-1}+s^*(\overline{X}_{n-1}) \wedge \Theta_n] \cap \mathbb{R}_+} e^{-\alpha s} \\
& \times \big[\mathcal{X}W(X(s)) - \alpha W(X(s))\big] ds + I_{\{\Theta_n > s^*(\overline{X}_{n-1})\}} \Big(e^{-\alpha T_n} W(X(T_n-)) \\
& - e^{-\alpha(T_{n-1}+s^*(\overline{X}_{n-1}))} W(X(T_{n-1} + s^*(\overline{X}_{n-1})-))\Big) \Bigg\} \Bigg] \\
& + \mathbb{E}^u_{x_0}\Bigg[\int_{([0,T_N] \cap \mathbb{R}_+) \times \mathbf{X}} e^{-\alpha s} \Big[W(z) - W(X(s-))\Big] \nu(ds, \overline{\mathbf{X}}, \mathbf{A}^i_\Delta, dz)\Bigg].
\end{aligned}
$$

and by using Lemma 2.5 we get the result. □

The following corollary will be useful in the proof of Proposition 5.6.

**Corollary 4.4** *In the framework of Lemma 4.3, we have for any $N \in \mathbb{N}^*$*

$$
\begin{aligned}
&\mathbb{E}^u_{x_0}\bigg[\int_{[0,T_N]\cap\mathbb{R}_+} e^{-\alpha s}\int_{\mathbf{A}^g(X(s))} C^g(X(s),a^g)\pi(da^g|s)ds\bigg] \qquad (9)\\
&+\int_{\mathbf{A}^i_\Delta(x_0)} C^i(x_0,a^i)u_0(da^i|x_0)\\
&+\mathbb{E}^u_{x_0}\bigg[\int_{([0,T_N]\cap\mathbb{R}_+)\times\overline{\mathbf{X}}\times\mathbf{A}^i_\Delta} e^{-\alpha s}C^i(x,a^i)\mu(ds,dx,da^i,\mathbf{X})\bigg]\\
&=\int_{\mathbf{A}^i_\Delta(x_0)}\int_{\mathbf{X}} W(y)Q(dy|x_0,a^i)u_0(da^i|x_0)+\int_{\mathbf{A}^i_\Delta(x_0)} C^i(x_0,a^i)u_0(da^i|x_0)\\
&+\mathbb{E}^u_{x_0}\bigg[\sum_{n=1}^N\bigg\{\int_{]T_{n-1},T_{n-1}+s^*(\overline{X}_{n-1})\wedge\Theta_n]\cap\mathbb{R}_+} e^{-\alpha s}\bigg\{\big[\mathcal{X}W(X(s))-\alpha W(X(s))\big]\\
&+\int_{\mathbf{A}^g(X(s))}\bigg[C^g(X(s),a^g)+\int_{\mathbf{X}}\int_{\mathbf{A}^i_\Delta(y)}\int_{\mathbf{X}}[C^i(y,a^i)+W(z)-W(X(s))]\\
&\times Q(dz|y,a^i)\gamma^0(da^i|y,s)Q(dy|X(s),a^g)\lambda(X(s),a^g)\bigg]\pi(da^g|s)\bigg\}ds\\
&+I_{\{\Theta_n>s^*(\overline{X}_{n-1})\}}\bigg(e^{-\alpha T_n}W(X(T_n-))-e^{-\alpha(T_{n-1}+s^*(\overline{X}_{n-1}))}W(X(T_{n-1}+s^*(\overline{X}_{n-1})-))\\
&+\int_{]T_{n-1}+s^*(\overline{X}_{n-1}),T_n]\cap\mathbb{R}_+} e^{-\alpha s}\int_{\mathbf{A}^g(X(s))}\bigg[C^g(X(s),a^g)+\int_{\mathbf{X}}\int_{\mathbf{A}^i_\Delta(y)}\int_{\mathbf{X}}[C^i(y,a^i)+W(z)-W(X(s))]\\
&\times Q(dz|y,a^i)\gamma^0(da^i|y,s)Q(dy|X(s),a^g)\lambda(X(s),a^g)\bigg]\pi(da^g|s)ds\bigg)\bigg\}\bigg]\\
&+\mathbb{E}^u_{x_0}\bigg[\sum_{n=1}^N\int_{]T_{n-1},T_n]\cap\mathbb{R}_+} e^{-\alpha s}\int_{\mathbf{A}^i(X(s-))}\bigg\{C^i(X(s-),a^i)+\int_{\mathbf{X}}\Big[W(z)-W(X(s-))\Big]\bigg\}\\
&\times Q(dz|X(s-),a^i)\gamma^1_n(da^i|H_{n-1},X(s-))\frac{\psi_n(ds-T_{n-1}|H_{n-1})}{\psi_n([s-T_{n-1},\infty]|H_{n-1})}\bigg]-\mathbb{E}^u_{x_0}\left[e^{-\alpha T_N}W(X(T_N))\right].
\end{aligned}
$$

**Proof:** This is a straightforward consequence of lemmas 2.5 and 4.3; see also formula (7). □

**Remark 4.5** *If $\psi_n(dt|h_{n-1})=\delta_{s^*(x_{n-1})}(dt)$, where the measurable function $s^*:\overline{\mathbf{X}}\to\overline{\mathbb{R}}_+$ satisfies $s^*(x)\le t^*(x)$, then (8) and (9) hold for $W\in\mathbb{A}_{s^*}(\overline{\mathbf{X}})$ provided $\mathcal{X}W\in\mathbb{B}(\overline{\mathbf{X}})$.*

# 5 Dynamic programming approach

In this section we present our main results. We provide in Theorem 5.5 sufficient conditions based on the three local characteristics of the process $\phi$, $\lambda$, $Q$, and the semi-continuity properties of the set valued action space, for the existence of a solution for an integro-differential HJB optimality equation associated with the problem as well as conditions for the existence of an optimal selector for this equation. In the sequel this result is used in Proposition 5.6 to show that the solution of the integro-differential HJB optimality equation is in fact unique and coincides with the optimal value of the criterion. Moreover, the optimal selector derived in Theorem 5.5 yields an optimal stationary non-randomized strategy for the problem. But before showing these results we need three auxiliary results presented in Lemmas 5.2, 5.3, and 5.4. First let us introduce the following definition.

**Definition 5.1** *Let $\beta > 0$ be a real constant, $s^* : \overline{\mathbf{X}} \to \overline{\mathbb{R}}_+$ be a measurable function satisfying $s^*(x) \leq t^*(x)$, and $F$ and $G$ be bounded $\mathbb{R}$-valued measurable functions on $\overline{\mathbf{X}}$. Then $\mathbf{\Xi}_{s^*}$ is the subset of $\overline{\mathbf{X}}$ defined by $\mathbf{\Xi}_{s^*} = \{x \in \overline{\mathbf{X}} : s^*(x) = 0\}$. Moreover, we say that $\beta, F, G, s^*$ are consistent if, for any $x \in \overline{\mathbf{X}}$, in case there is $t < s^*(x)$ such that $t + s^*(\phi(x,t)) \neq s^*(x)$, the following equality holds:*

$$\begin{aligned} & \int_{[0,s^*(\phi(x,t))]\cap\mathbb{R}_+} e^{-\beta s} F(\phi(\phi(x,t),s))ds + e^{-\beta s^*(\phi(x,t))} G(\phi(\phi(x,t),s^*(\phi(x,t)))) \\ = & \int_{[0,s^*(x)-t]\cap\mathbb{R}_+} e^{-\beta s} F(\phi(\phi(x,t),s))ds + e^{-\beta(s^*(x)-t)} G(\phi(\phi(x,t),s^*(x)-t)). \end{aligned} \tag{10}$$

**Lemma 5.2** *Suppose $s^* : \overline{\mathbf{X}} \to \overline{\mathbb{R}}_+$ is a measurable function satisfying $s^*(x) \leq t^*(x)$. Consider a bounded $\mathbb{R}$-valued measurable function $F$ (respectively, $G$) defined on $\overline{\mathbf{X}}$ (respectively, $\mathbf{\Xi}_{s^*}$) and a real number $\beta > 0$ such that $\beta, F, G, s^*$ are consistent.*

*Then the mapping $V$ defined on $\overline{\mathbf{X}}$ by*

$$V(x) = \int_{[0,s^*(x)]\cap\mathbb{R}_+} e^{-\beta s} F(\phi(x,s))ds + e^{-\beta s^*(x)} G(\phi(x,s^*(x)))$$

*belongs to $\mathbb{A}_{s^*}(\overline{\mathbf{X}})$ and $V(z) = G(z)$ for any $z \in \mathbf{\Xi}_{s^*}$. Moreover, the function $\mathcal{X}V \in \mathbb{B}(\overline{\mathbf{X}})$ can be taken of the form*

$$\mathcal{X}V(x) = \beta V(x) - F(x), \qquad x \in \overline{\mathbf{X}}.$$

**Proof:** The proof is similar to the proof of Lemma 5.1 in [4] and based on the representation

$$V(\phi(x,t)) = e^{\beta t} \left\{ \int_{[t,s^*(x)]\cap\mathbb{R}_+} e^{-\beta s} F(\phi(x,s))ds + e^{-\beta s^*(x)} G(\phi(x,s^*(x))) \right\},$$

which is valid for any $x \in \overline{\mathbf{X}}$, $t \in [0, s^*(x)] \cap \mathbb{R}_+$ due to the fact that $\beta, F, G, s^*$ are consistent. As a consequence, we have

$$\int_{[0,t]} [\beta V(\phi(x,s)) - F(\phi(x,s))]ds = V(\phi(x,t)) - V(x)$$

for any $x \in \overline{\mathbf{X}}$ and $t \in [0, s^*(x)] \cap \mathbb{R}_+$ showing the last part of the result. □

For a fixed number $K$, let us introduce for any $V \in \mathbb{B}(\overline{\mathbf{X}})$ the real-valued function $\mathfrak{R}V$ defined on $\overline{\mathbf{X}}$ by

$$\mathfrak{R}V(x) \;=\; \inf_{a^g \in \mathbf{A}^g(x)} \Big\{ C^g(x,a^g) + qV(x,a^g) + KV(x) \Big\}, \tag{11}$$

where $q$ is defined in equation (1). For notational convenience, let us denote the real-valued function $\mathfrak{T}V$ defined on $\overline{\mathbf{X}}^i$ by

$$\mathfrak{T}V(z) \;=\; \inf_{a^i \in \mathbf{A}^i(z)} \Big\{ C^i(z,a^i) + QV(z,a^i) \Big\}, \tag{12}$$

for any $V \in \mathbb{B}(\overline{\mathbf{X}})$.

Finally,

$$\mathfrak{B}V(y) = \inf_{t \in \mathbb{T}^i_y} \left\{ \int_{[0,t]\cap\mathbb{R}_+} e^{-(K+\alpha)u} \mathfrak{R}V(\phi(y,u))du + e^{-(K+\alpha)t} \mathfrak{T}V(\phi(y,t)) \right\} \tag{13}$$

is a function on $\overline{\mathbf{X}}$.

**Lemma 5.3** *Suppose Assumptions (A1), B and C are satisfied and $K \geq K_1$. If $V \in \mathbb{L}(\overline{\mathbf{X}})$ then $\mathfrak{R}V \in \mathbb{L}(\overline{\mathbf{X}})$, $\mathfrak{T}V \in \mathbb{L}(\overline{\mathbf{X}}^i)$, and $\mathfrak{B}V \in \mathbb{L}(\overline{\mathbf{X}})$. Moreover, there is a measurable map $s^*$ from $\overline{\mathbf{X}}$ to $\overline{\mathbb{R}}_+$ providing the infimum in (13) for any $y \in \overline{\mathbf{X}}$. The set $(K+\alpha), \mathfrak{R}V, \mathfrak{T}V, s^*$ is consistent. (One can consider an arbitrary bounded measurable extension of $\mathfrak{T}V$ to the whole space $\overline{\mathbf{X}}$.)*

**Proof:** Consider $V \in \mathbb{L}(\overline{\mathbf{X}})$. By using hypotheses (B2)-(B3) and the fact that $\lambda$ is bounded by $K_1$ on $\mathbf{K}^g$, we obtain that $qV + KV \in \mathbb{L}(\mathbf{K}^g)$, and so, $C^g + qV + KV \in \mathbb{L}(\mathbf{K}^g)$ by Assumption (C3). Therefore, combining Lemma 17.30 in [1] with Assumptions (B1) and (C1), it yields that $\mathfrak{R}V \in \mathbb{L}(\overline{\mathbf{X}})$. By using the same arguments, it can be shown that $\mathfrak{T}V \in \mathbb{L}(\overline{\mathbf{X}}^i)$.

Further, the function of $(y,t)$ given by $e^{-(K+\alpha)t}\mathfrak{T}V(\phi(y,t))$ is bounded and lower semi-continuous on $\mathbb{Y} = \{(y,t) : y \in \overline{\mathbf{X}}, t \in \mathbb{T}^i_y\}$. (With some abuse of notation, we accept that $e^{-(K+\alpha)t}\mathfrak{T}V(\phi(y,t)) = 0$ if $t = \infty$.) To show this, let $f$ be a bounded continuous function on $\overline{\mathbf{X}}^i$. If $(y_n, t_n) \to (y,t)$ in $\mathbb{Y}$ as $n \to \infty$, then, in case $t < \infty$, $\phi(y_n,t_n) \to \phi(y,t)$ and $e^{-(K+\alpha)t_n} f(\phi(y_n,t_n)) \to e^{-(K+\alpha)t} f(\phi(y,t))$. In case $t = \infty$, $e^{-(K+\alpha)t_n} f(\phi(y_n,t_n)) \to 0$ as well. To complete this step of the proof, remember that any function is bounded below and lower semi-continuous if and only if it is the limit of an increasing sequence of bounded continuous functions.

Similarly, for any bounded continuous function $f$ on $\overline{\mathbf{X}}^i$, the function $\int_{[0,t]\cap\mathbb{R}_+} e^{-(K+\alpha)u} f(\phi(y,u))du$ is continuous on $\mathbb{Y}$, and hence the function $\int_{[0,t]\cap\mathbb{R}_+} e^{-(K+\alpha)u}\mathfrak{R}V(\phi(y,u))du$ is bounded and lower semi-continuous on $\mathbb{Y}$.

Since the set $\mathbb{Y}$ is closed in $\overline{\mathbf{X}} \times \overline{\mathbb{R}}_+$ and $\overline{\mathbb{R}}_+$ is compact, we conclude by Proposition 7.33 in [2] that $\mathfrak{B}V \in \mathbb{L}(\overline{\mathbf{X}})$ and the required mapping $s^*$ does exist.

To prove the consistency, suppose, for some $x \in \overline{\mathbf{X}}$ and $t < s^*(x)$, $t + s^*(\phi(x,t)) \neq s^*(x)$ and equality (10) is violated. If the lefthand side in (10) is smaller than the righthand side then

$$\begin{aligned} & e^{(K+\alpha)t} \int_{[t,t+s^*(\phi(x,t))]\cap\mathbb{R}_+} e^{-(K+\alpha)u}\mathfrak{R}V(\phi(x,u))du + e^{-(K+\alpha)s^*(\phi(x,t))}\mathfrak{T}V(\phi(x,s^*(\phi(x,t)))) \\ < \; & e^{(K+\alpha)t}\Big[\int_{[t,s^*(x)]\cap\mathbb{R}_+} e^{-(K+\alpha)u}\mathfrak{R}V(\phi(x,u))du + e^{-(K+\alpha)s^*(x)}\mathfrak{T}V(\phi(x,s^*(x)))\Big]. \end{aligned}$$

Now, multiplying the both terms of the previous equation by $e^{-(K+\alpha)t}$ and adding the following expression $\int_{[0,t]\cap\mathbb{R}_+} e^{-(K+\alpha)u}\mathfrak{R}V(\phi(x,u))du$, we get

$$\begin{aligned} & \int_{[0,t+s^*(\phi(x,t))]\cap\mathbb{R}_+} e^{-(K+\alpha)u}\mathfrak{R}V(\phi(x,u))du + e^{-(K+\alpha)(t+s^*(\phi(x,t)))}\mathfrak{T}V(\phi(x,t+s^*(\phi(x,t)))) \\ < \; & \int_{[0,s^*(x)]\cap\mathbb{R}_+} e^{-(K+\alpha)u}\mathfrak{R}V(\phi(x,u))du + e^{-(K+\alpha)s^*(x)}\mathfrak{T}V(\phi(x,s^*(x))). \end{aligned}$$

Now observe that $t + s^*(\phi(x,t)) \in \mathbb{T}^i_x$ since $s^*(x) \in \mathbb{T}^i_x$. Therefore, $s^*(x)$ does not provide the infimum in (13) leading to a contradiction.

If the righthand side in (10) is smaller, then, similarly, $s^*(\phi(x,t))$ does not provide the infimum in (13) at $y = \phi(x,t)$. $\square$

Suppose Assumptions A are satisfied and $K \geq K_1, K_3$. Let us introduce the constant

$$K_A \quad = \quad \frac{K(1+K_4)(1-e^{-(K+\alpha)\varepsilon}) + (K+\alpha)K_4 e^{-(K+\alpha)\varepsilon}}{\alpha(1-e^{-(K+\alpha)\varepsilon})}. \tag{14}$$

**Lemma 5.4** *Suppose that Assumptions A, B and C are satisfied and $K \geq K_1$. Consider an arbitrary function $V \in \mathbb{L}(\overline{\mathbf{X}})$.*

*(i) Let $s^* \leq t^*$ be the mapping from Lemma 5.3 and let $\mathbf{\Xi}_{s^*} = \{x \in \overline{\mathbf{X}} : s^*(x) = 0\}$. Then $W = \mathfrak{B}V \in \mathbb{A}_{s^*}(\overline{\mathbf{X}})$, the set $(K+\alpha), \mathfrak{R}V, \mathfrak{T}V, s^*$ is consistent, there exists a function $\mathcal{X}W \in \mathbb{B}(\overline{\mathbf{X}})$ satisfying*

$$-(K+\alpha)W(x) + \mathcal{X}W(x) = -\mathfrak{R}V(x)$$

*for any $x \in \overline{\mathbf{X}}$ and $W(z) = \mathfrak{T}V(z)$ for any $z \in \mathbf{\Xi}_{s^*}$.*

*(ii) If $|V(y)| \leq K_A I_{A_\varepsilon}(y) + (K_A + K_4) I_{A_\varepsilon^c}(y)$ for any $y \in \overline{\mathbf{X}}$ then $|\mathfrak{B}V(y)| \leq K_A I_{A_\varepsilon}(y) + (K_A + K_4) I_{A_\varepsilon^c}(y)$.*

**Proof:** (i) According to Lemma 5.3, the equality (13) can be rewritten as

$$\mathfrak{B}V(y) = \int_{[0,s^*(y)]\cap\mathbb{R}_+} e^{-(K+\alpha)u} \mathfrak{R}V(\phi(y,u))du + e^{-(K+\alpha)s^*(y)} \mathfrak{T}V(\phi(y, s^*(y))).$$

Now all the assertions follow from lemmas 5.2 and 5.3.

(ii) Let us first introduce the function $\widehat{\mathfrak{B}}V$ defined on $\overline{\mathbf{X}}$ by

$$\widehat{\mathfrak{B}}V(y) = \int_{[0,t^*(y)]\cap\mathbb{R}_+} e^{-(K+\alpha)u} \mathfrak{R}V(\phi(y,u))du + e^{-(K+\alpha)t^*(y)} \mathfrak{T}V(\phi(y, t^*(y))).$$

By using similar arguments as those for the proof of Lemma 5.4 in [4], it can be shown that $|\widehat{\mathfrak{B}}V(y)| \leq K_A I_{A_\varepsilon}(y) + (K_A + K_4) I_{A_\varepsilon^c}(y)$ implying that $\mathfrak{B}V(y) \leq K_A I_{A_\varepsilon}(y) + (K_A + K_4) I_{A_\varepsilon^c}(y)$ for any $y \in \overline{\mathbf{X}}$. Consider now the function $\widetilde{\mathfrak{B}}V$ defined on $\mathbf{X}$ by

$$\widetilde{\mathfrak{B}}V(y) = \int_{[0,t]\cap\mathbb{R}_+} e^{-(K+\alpha)u} \mathfrak{R}V(\phi(y,u))du + e^{-(K+\alpha)t} \mathfrak{T}V(\phi(y,t)),$$

with fixed $t \in \mathbb{T}_y^i$, such that $t < t^*(y)$. For $(z,b) \in \mathbf{K}^i$ with $z \notin \mathbf{\Xi}$, we have $C^i(z,b) > K_2$ and $QV(z,b) \geq -K_A - K_4$. Therefore,

$$\mathfrak{T}V(z) \geq K_2 - K_A - K_4 \geq -K_A - K_4.$$

When $x \in A_\varepsilon^c \cap \mathbf{X}$ and by using similar arguments as in the proof of Lemma 5.4 in [4], we can show that

$$\widetilde{\mathfrak{B}}V(x) \geq -K_A - K_4.$$

When $x \in A_\varepsilon$ and $t < t^*(x) < \infty$, we obtain by definition of $\widetilde{\mathfrak{B}}V$

$$\widetilde{\mathfrak{B}}V(x) \geq -\frac{K(1+K_A+K_4)}{K+\alpha}\left(1 - e^{-(K+\alpha)t}\right) + (K_2 - K_4 - K_A)e^{-(K+\alpha)t}.$$

However, observe that $K_2 - K_A - K_4 \geq -K_A$ and $-\frac{K(1+K_A+K_4)}{K+\alpha} \geq -K_A$ and so, $\widetilde{\mathfrak{B}}V(x) \geq -K_A$.

In conclusion, we have proved that $\widetilde{\mathfrak{B}}V(y) \geq -K_A I_{A_\varepsilon}(y) - (K_A + K_4) I_{A_\varepsilon^c}(y)$ showing $\mathfrak{B}V(y) \geq -K_A I_{A_\varepsilon}(y) - (K_A + K_4) I_{A_\varepsilon^c}(y)$ for $y \in \mathbf{X}$.

Finally, we have shown that $|\mathfrak{B}V(x)| \leq K_A I_{A_\varepsilon}(x) + (K_A + K_4) I_{A_\varepsilon^c}(x)$ for any $x \in \overline{\mathbf{X}}$ completing the proof of the item (ii). $\square$

The next theorem, provides sufficient conditions for the existence of a solution for the HJB equation associated with the optimization problem as well as conditions for the existence of an optimal selector for this equation.

**Theorem 5.5** *Suppose assumptions A, B and C are satisfied and $K \geq K_1$. Then equation $\mathfrak{B}W = W$ has a solution $W \in \mathbb{L}(\overline{\mathbf{X}})$ and the following assertions hold.*

*(i) There is a measurable map $s^* : \overline{\mathbf{X}} \to \overline{\mathbb{R}}_+$ providing the infimum in equation*

$$\mathfrak{B}W(y) = \inf_{t \in \mathbb{T}^i_y} \left\{ \int_{[0,t] \cap \mathbb{R}_+} e^{-(K+\alpha)u} \mathfrak{R}W(\phi(y,u))du + e^{-(K+\alpha)t} \mathfrak{T}W(\phi(y,t)) \right\}$$

*for any $y \in \overline{\mathbf{X}}$. Below*

$$\Xi_{s^*} = \{x \in \overline{\mathbf{X}} : s^*(x) = 0\} \subset \overline{\mathbf{X}}^i.$$

*The set $(K+\alpha), \mathfrak{R}W, \mathfrak{T}W, s^*$ is consistent.*

*(ii) $W \in \mathbb{A}_{s^*}(\overline{\mathbf{X}})$ and there exists function $\mathcal{X}W \in \mathbb{B}(\overline{\mathbf{X}})$ satisfying equation*

$$\begin{aligned} &-\alpha W(x) + \mathcal{X}W(x) + \inf_{a^g \in \mathbf{A}^g(x)} \bigg\{ C^g(x,a^g) + \int_{\mathbf{X}} \inf_{a^i \in \mathbf{A}^i_\Delta(y)} \{C^i(y,a^i) + QW(y,a^i)\} \\ &\times Q(dy|x,a^g)\, \lambda(x,a^g) - W(x)\lambda(x,a^g) \bigg\} = 0 \end{aligned} \tag{15}$$

*for any $x \in \overline{\mathbf{X}}$. For $z \in \Xi_{s^*}$,*

$$W(z) = \inf_{a^i \in A^i(z)} \Big\{ C^i(z,a^i) + QW(z,a^i) \Big\}. \tag{16}$$

*(iii) There are measurable mappings $\varphi^g : \overline{\mathbf{X}} \setminus \Xi_{s^*} \to \mathbf{A}^g$ and $\varphi^i : \Xi_{s^*} \to \mathbf{A}^i$ such that $\varphi^g(x) \in \mathbf{A}^g(x)$ (respectively, $\varphi^i(z) \in \mathbf{A}^i(z)$) provides the infimum in (15) with respect to $\mathbf{A}^g(x)$ (respectively, the infimum in (16) with respect to $\mathbf{A}^i(z)$). The map $\varphi^i$ provides also the infimum in*

$$\inf_{a^i \in \mathbf{A}^i_\Delta(y)} \{C^i(y,a^i) + QW(y,a^i)\} = W(y), \tag{17}$$

*if $y \in \Xi_{s^*}$; in case $y \in \overline{\mathbf{X}} \setminus \Xi_{s^*}$*

$$\inf_{a^i \in \mathbf{A}^i_\Delta(y)} \{C^i(y,a^i) + QW(y,a^i)\} = C^i(y,\Delta) + QW(y,\Delta) = W(y). \tag{18}$$

*(iv) For any $x \in \mathbf{X}$, for any $0 \leq t \leq T \leq t^*(x)$,*

$$\begin{aligned} &e^{-\alpha(T-t)} W(\phi(x,T)) - W(\phi(x,t)) \\ &+ \int_{]t,T] \cap \mathbb{R}_+} e^{-\alpha(s-t)} \inf_{a^g \in \mathbf{A}^g(\phi(x,s))} \{C^g(\phi(x,s),a^g) + qW(\phi(x,s),a^g)\}\, ds \geq 0. \end{aligned} \tag{19}$$

**Proof:** By Lemma 5.3, one can define recursively the sequence of functions $\{W_i\}_{i \in \mathbb{N}}$ in $\mathbb{L}(\overline{\mathbf{X}})$ as follows $W_{i+1}(y) = \mathfrak{B}W_i(y)$, for $i \in \mathbb{N}$ and $W_0(y) = -K_A I_{A_{\varepsilon_1}}(y) - (K_A + K_4) I_{A^c_{\varepsilon_1}}(y)$ for any $y \in \overline{\mathbf{X}}$. By using Lemma 5.4 (ii) and the definition of $W_0$, we obtain that $W_1(y) \geq W_0(y)$ for any $y \in \overline{\mathbf{X}}$. Now, observe that the operator $\mathfrak{B}$ is monotone, that is, $V_1 \leq V_2$ implies $\mathfrak{B}V_1 \leq \mathfrak{B}V_2$. Consequently, it can be shown by induction that the sequence $\{W_i\}_{i \in \mathbb{N}}$ is increasing. Using again Lemma 5.4 (ii), this sequence is uniformly bounded, that is, $\sup_{y \in \overline{\mathbf{X}}} |W_i(y)| \leq K_A + K_4$ for any $i \in \mathbb{N}$. As a result, $\{W_i\}_{i \in \mathbb{N}}$ converges to a mapping $W \in \mathbb{B}(\overline{\mathbf{X}})$. Since $\{W_i\}_{i \in \mathbb{N}}$ is an increasing sequence of lower semi-continuous functions, $W \in \mathbb{L}(\overline{\mathbf{X}})$, $qW_i + KW_i \in \mathbb{L}(\mathbf{K}^g)$, $C^g + qW_i + KW_i \in \mathbb{L}(\mathbf{K}^g)$ and $qW + KW \in \mathbb{L}(\mathbf{K}^g)$, $C^g + qW + KW \in \mathbb{L}(\mathbf{K}^g)$ by Assumption (C3). Therefore, combining

Assumptions (B1) and (C1) and Lemma 2.1 in [14], it follows that $\lim_{i\to\infty} \mathfrak{R}W_i(x) = \mathfrak{R}W(x)$, for any $x \in \overline{\mathbf{X}}$ and $\lim_{i\to\infty} \mathfrak{T}W_i(z) = \mathfrak{T}W(z)$ for any $z \in \overline{\mathbf{X}}^i$. These both convergences are monotone (increasing). By using the monotone convergence theorem, it implies that for any $y \in \overline{\mathbf{X}}$, $t \in \mathbb{T}_y^i$:

$$\begin{aligned} & \lim_{i\to\infty} \left\{ \int_{[0,t]\cap\mathbb{R}_+} e^{-(K+\alpha)t} \mathfrak{R}W_i(\phi(y,t))dt + e^{-(K+\alpha)t}\mathfrak{T}W_i(\phi(y,t)) \right\} \\ = & \int_{[0,t]\cap\mathbb{R}_+} e^{-(K+\alpha)t}\mathfrak{R}W(\phi(y,t))dt + e^{-(K+\alpha)t}\mathfrak{T}W(\phi(y,t)) = \mathfrak{B}W(y). \end{aligned}$$

Finally again using Lemma 2.1 in [14], Assumptions (B6), (C2) and the previous equation, we obtain

$$\begin{aligned} W(y) &= \lim_{i\to\infty} \inf_{t\in\mathbb{T}_y^i} \left\{ \int_{[0,t]\cap\mathbb{R}_+} e^{-(K+\alpha)t}\mathfrak{R}W_i(\phi(y,t))dt + e^{-(K+\alpha)t}\mathfrak{T}W_i(\phi(y,t)) \right\} \quad (20) \\ &= \inf_{t\in\mathbb{T}_y^i} \left\{ \int_{[0,t]\cap\mathbb{R}_+} e^{-(K+\alpha)t}\mathfrak{R}W(\phi(y,t))dt + e^{-(K+\alpha)t}\mathfrak{T}W(\phi(y,t)) \right\} = \mathfrak{B}W(y) \end{aligned}$$

for any $y \in \overline{\mathbf{X}}$.

Assertion (i) follows directly from Lemma 5.3.

For assertion (ii), note that Lemma 5.4 implies that $W \in \mathbb{A}_{s^*}(\overline{\mathbf{X}})$ and there exists a function $\mathcal{X}W \in \mathbb{B}(\overline{\mathbf{X}})$ such that

$$-(K+\alpha)W(x) + \mathcal{X}W(x) = -\mathfrak{R}W(x) = -\inf_{a^g\in\mathbf{A}^g(x)} \Big\{ C^g(x,a^g) + qW(x,a^g) + KW(x) \Big\} \quad (21)$$

for all $x \in \overline{\mathbf{X}}$. Clearly, we have $W(z) = \mathfrak{T}W(z)$ on $\mathbf{\Xi}_{s^*}$ giving (16). Equation (15) follows from (21) and will be justified later.

(iii) Proposition D.5 in [12] implies the existence of the mappings $\varphi^g$ and $\varphi^i$ providing the infima in (16) and in (21). The mapping $\varphi^i$ provides also the infimum in (17) because $C^i(y,\Delta) + QW(y,\Delta) = W(y)$. Now, if we substitute (17) into (21) and cancel $KW(x)$, we get equation (15).

Equation (18) obviously holds if $\mathbf{A}^i(y) = \emptyset$. Suppose $\mathbf{A}^i(y) \neq \emptyset$, that is, $0 \in \mathbb{T}_y^i$ and $W(y) = \mathfrak{B}W(y) \leq \mathfrak{T}W(y)$. Hence $W(y) \leq \inf_{a^i\in\mathbf{A}_\Delta^i(y)}\{C^i(y,a^i) + QW(y,a^i)\}$. The reverse inequality $W(y) \geq \inf_{a^i\in\mathbf{A}_\Delta^i(y)}\{C^i(y,a^i) + QW(y,a^i)\}$ is obvious.

(iv) Denote for brevity $\mathfrak{F}(x) = \inf_{a^g\in\mathbf{A}^g(x)}\{C^g(x,a^g) + qW(x,a^g)\}$. Consider the functions

$$g_t^T(x) = \int_{]t,T]\cap\mathbb{R}_+} e^{-\alpha(s-t)}\mathfrak{F}(\phi(x,s))ds + e^{-\alpha(T-t)}W(\phi(x,T)) \quad (22)$$

and

$$V_t^T(x) = \int_{]t,T]\cap\mathbb{R}_+} e^{-(K+\alpha)(s-t)}[\mathfrak{F}(\phi(x,s)) + KW(\phi(x,s))]ds + e^{-(K+\alpha)(T-t)}W(\phi(x,T)).$$

Note that

$$\begin{aligned} W(\phi(x,T)) &= \mathfrak{B}W(\phi(x,T)) = \inf_{\tilde{t}\in\mathbb{T}^i_{\phi(x,T)}} \left\{ \int_{[0,\tilde{t}]\cap\mathbb{R}_+} e^{-(K+\alpha)u}\mathfrak{R}W(\phi(\phi(x,T),u))du \right. \\ & \left. \quad + e^{-(K+\alpha)\tilde{t}}\mathfrak{T}W(\phi(\phi(x,T),\tilde{t})) \right\}. \end{aligned}$$

Now, recalling that $\mathfrak{R}W(x) = \mathfrak{F}(x) + KW(x)$ and by using the previous equation we get

$$\begin{aligned} V_t^T(x) &= \inf_{\tilde{t}\in\mathbb{T}^i_{\phi(x,T)}} \left\{ e^{-(K+\alpha)(T-t)} \int_{[0,\tilde{t}]\cap\mathbb{R}_+} e^{-(K+\alpha)u} \mathfrak{R}W(\phi(\phi(x,T),u))du \right. \\ &\quad \left. + \int_{]t,T]\cap\mathbb{R}_+} e^{-(K+\alpha)(s-t)} \mathfrak{R}W(\phi(x,s))ds + e^{-(K+\alpha)(T-t+\tilde{t})} \mathfrak{T}W(\phi(x,T+\tilde{t})) \right\}. \end{aligned}$$

By using a change of variable, we obtain

$$V_t^T(x) = \inf_{\tilde{t}\in\mathbb{T}^i_{\phi(x,T)}} \left\{ \int_{]t,T+\tilde{t}]\cap\mathbb{R}_+} e^{-(K+\alpha)(s-t)} \mathfrak{R}W(\phi(x,s))ds + e^{-(K+\alpha)(T-t+\tilde{t})} \mathfrak{T}W(\phi(x,T+\tilde{t})) \right\}.$$

Now, observe that $\tilde{t} \in \mathbb{T}^i_{\phi(x,T)}$ is equivalent to $T+\tilde{t} \in \mathbb{T}^i_x$ and so,

$$V_t^T(x) = \inf_{\substack{\tilde{t}\geq 0 \\ T+\tilde{t}\in\mathbb{T}^i_x}} \left\{ \int_{]t,T+\tilde{t}]\cap\mathbb{R}_+} e^{-(K+\alpha)(s-t)} \mathfrak{R}W(\phi(x,s))ds + e^{-(K+\alpha)(T-t+\tilde{t})} \mathfrak{T}W(\phi(x,T+\tilde{t})) \right\}.$$

Using a change of variable ($r = T+\tilde{t}-t$, $z = \phi(x,t)$, $u = s-t$) we obtain

$$\begin{aligned} V_t^T(x) &= \inf_{\substack{r\geq T-t \\ r\in\mathbb{T}^i_{\phi(x,t)}}} \left\{ \int_{]0,r]\cap\mathbb{R}_+} e^{-(K+\alpha)u} \mathfrak{R}W(\phi(x,u+t))du + e^{-(K+\alpha)r} \mathfrak{T}W(\phi(x,r+t)) \right\} \\ &\geq \inf_{r\in\mathbb{T}^i_z} \left\{ \int_{]0,r]\cap\mathbb{R}_+} e^{-(K+\alpha)u} \mathfrak{R}W(\phi(z,u))du + e^{-(K+\alpha)r} \mathfrak{T}W(\phi(z,r) \right\} \\ &= W(z) = W(\phi(x,t)). \end{aligned} \tag{23}$$

The function $g_t^T$ can be equivalently rewritten as

$$g_t^T(x) = \int_{]t,T]\cap\mathbb{R}_+} e^{-(K+\alpha)(s-t)} [\mathfrak{F}(\phi(x,s)) + Kg_s^T(x)]ds + e^{-(K+\alpha)(T-t)} W(\phi(x,T)) \tag{24}$$

because the both functions (22) and (24) of $t \in [0,T]$ satisfy the boundary condition $g_T^T(x) = W(\phi(x,T))$ and the following ordinary differential equation:

$$\frac{dg_t^T(x)}{dt} = \alpha g_t^T(x) - \mathfrak{F}(\phi(x,t)).$$

Finally, for any $t \in [0,T]$, from (23) and (24) we have

$$W(\phi(x,t)) - g_t^T(x) \leq V_t^T(x) - g_t^T(x) = K \int_{]t,T]\cap\mathbb{R}_+} e^{-(K+\alpha)(u-t)} [W(\phi(x,u)) - g_u^T(x)]du,$$

and the required formula (19) follows from the Gronwall inequality: $W(\phi(x,t)) - g_t^T(x) \leq 0$. □

It will shortly be clear that the optimal strategy prescribes to apply impulsive actions if and only if the current state of the controlled process $X(t)$ belongs to $\boldsymbol{\Xi}_{s^*}$. In general, this can lead to a sequence of simultaneous interventions, and such case is not within the scope of the current work. At the same time, it is clear that this can never happen under the conditions enlisted in Remark 2.2.

The following proposition shows that the solution of the HJB equation is in fact unique and coincides with the optimal value for the optimal control problem. Moreover this result provides the existence of an optimal stationary non-randomized strategy.

**Proposition 5.6** *Suppose all the conditions of Theorem 5.5 are satisfied. Let $s^*$, $\varphi^g$ and $\varphi^i$ be the corresponding mappings, extend $\varphi^g$ to $\overline{\mathbf{X}}$ in an arbitrary measurable way, put $\varphi^i(x) = \Delta$ for $x \in \overline{\mathbf{X}} \setminus \mathbf{\Xi}_{s^*}$, and consider the stationary non-randomized control strategy $u^*$ defined as follows*

- $u_0(\cdot|y) = \delta_{\varphi^i(y)}(\cdot)$;
- $\psi_n(\cdot|h_{n-1}) = \delta_{s^*(\overline{x}_{n-1})}(\cdot)$;
- $\gamma_n^0(\cdot|h_{n-1}, y) = \gamma_n^1(\cdot|h_{n-1}, y) = \delta_{\varphi^i(y)}(\cdot)$.

*Assume, $u^* \in \mathcal{U}$, i.e. this strategy does not lead to sequences of simultaneous interventions. Then the following assertion holds.*

*Strategy $u^*$ is uniformly optimal, the function $W$ is the unique solution to the HJB equation $\mathfrak{B}W = W$ in the class $\mathbb{L}(\overline{\mathbf{X}})$ and coincides with $\inf_{u\in\mathcal{U}} \mathcal{V}(u, x)$.*

**Proof:** Equations (9), (15), (16), (17), (18), and (19) imply that $\forall u \in \mathcal{U}$, $\forall x_0 \in \mathbf{X}$, $\forall n \in \mathbb{N}^*$

$$\begin{aligned}
&\int_{\mathbf{A}^i_\Delta(x_0)} \int_{\mathbf{X}} W(y) Q(dy|x_0, a^i) u_0(da^i|x_0) + \int_{\mathbf{A}^i_\Delta(x_0)} C^i(x_0, a^i) u_0(da^i|x_0) \\
&+\mathbb{E}^u_{x_0}\Bigg[\int_{]T_{n-1}, T_{n-1}+s^*(\overline{X}_{n-1}\wedge\Theta_n]\cap\mathbb{R}_+} e^{-\alpha s}\Bigg\{\big[\mathcal{X}W(X(s)) - \alpha W(X(s))\big] \\
&+\int_{\mathbf{A}^g(X(s))}\Bigg[C^g(X(s), a^g) + \int_{\mathbf{X}}\int_{\mathbf{A}^i_\Delta(y)}\int_{\mathbf{X}}[C^i(y, a^i) + W(z) - W(X(s))] \\
&\times Q(dz|y, a^i)\gamma^0(da^i|y, s) Q(dy|X(s), a^g)\lambda(X(s), a^g)\Bigg]\pi(da^g|s)\Bigg\}\,ds\Bigg] \\
&+\mathbb{E}^u_{x_0}\Bigg[I_{\{\Theta_n > s^*(\overline{X}_{n-1})\}}\Bigg(e^{-\alpha T_n} W(X(T_n-)) - e^{-\alpha(T_{n-1}+s^*(\overline{X}_{n-1}))} W(X(T_{n-1} + s^*(\overline{X}_{n-1})-)) \\
&+\int_{]T_{n-1}+s^*(\overline{X}_{n-1}), T_n]\cap\mathbb{R}_+} e^{-\alpha s}\int_{\mathbf{A}^g(X(s))}\Bigg[C^g(X(s), a^g) + \int_{\mathbf{X}}\int_{\mathbf{A}^i_\Delta(y)}\int_{\mathbf{X}}[C^i(y, a^i) + W(z) - W(X(s))] \\
&\times Q(dz|y, a^i)\gamma^0(da^i|y, s) Q(dy|X(s), a^g)\lambda(X(s), a^g)\Bigg]\pi(da^g|s)ds\Bigg)\Bigg] \\
&+\mathbb{E}^u_{x_0}\Bigg[\sum_{n=1}^{N}\int_{]T_{n-1}, T_n]\cap\mathbb{R}_+} e^{-\alpha s}\int_{\mathbf{A}^i(X(s-))}\Bigg\{C^i(X(s-), a^i) + \int_{\mathbf{X}}\Big[W(z) - W(X(s-))\Big]\Bigg\} \\
&\times Q(dz|X(s-), a^i)\gamma^1_n(da^i|H_{n-1}, X(s-))\frac{\psi_n(ds - T_{n-1}|H_{n-1})}{\psi_n([s - T_{n-1}, \infty]|H_{n-1})}\Bigg] \\
&\geq W(x_0) - \mathbb{E}^u_{x_0}\left[e^{-\alpha T_N} W(X(T_N))\right].
\end{aligned} \tag{25}$$

If $\mathcal{V}(u, x_0) < \infty$, then, due to Lemma 4.1, $\mathbb{P}^u_{x_0}(T_\infty < +\infty) = 0$ and hence $\mathcal{V}(u, x_0) \geq W(x_0)$. The last inequality is obvious if $\mathcal{V}(u, x_0) = \infty$. Now, for $u^*$, we have the equality in (25). Hence $\mathcal{V}(u^*, x_0) < \infty$ because $W$ is bounded, and, similarly to the presented above reasoning, $\mathcal{V}(u^*, x_0) = W(x_0)$.

Therefore, $\forall x_0 \in \mathbf{X}$ function $W(x_0) = \inf_{u\in\mathcal{U}} \mathcal{V}(u, x_0)$ is unique because, for any other solution $\widetilde{W}$ to the HJB equation $\mathfrak{B}W = W$ in the class $\mathbb{L}(\overline{\mathbf{X}})$, all the statements of Theorem 5.5 are valid; hence $\widetilde{W}(x_0) = \inf_{u\in\mathcal{U}} \mathcal{V}(u, x_0)$. □

# 6 Possible extensions

We investigated the time-homogeneous discounted model. But the developed theory can be also applied to the non-homogeneous process on the finite horizon $[0, T]$ when functions $\lambda(x, t, a)$, $Q(\cdot|x, t, a)$

and $C^{i,g}(x,t,a)$ depend on time $t$. Indeed, one can incorporate the $t$ variable in the state and consider the PDMP in the space $\mathbb{R}^{d+1}$, where the $(d+1)$-st component of the flow corresponds to the differential equation $\dot{t} = 1$ with the initial value $t(0) = 0$. The modification of the state space is obvious. For $t \in [0,T)$, $\mathbf{A}^i((x,t)) = \mathbf{A}^i(x)$, $\mathbf{A}^g((x,t)) = \mathbf{A}^g(x)$. With some additional efforts, the sets of feasible actions can be also made time-dependent. Any one point $x$ from the original boundary $\partial\mathbf{X}$ transforms to the segment $\{(x,t) : t \in [0,T)\}$. We also say that the set $\{(x,T),\ x \in \overline{\mathbf{X}}\}$ belongs to the boundary, where $\overline{\mathbf{X}}$ corresponds to the original model. For any point $(x,T)$ we put $\mathbf{A}^i((x,T)) = a^i_\infty$, $\mathbf{A}^g((x,T)) = a^g_\infty$ (artificial isolated points which are different in order to satisfy $\mathbf{A}^i \cap \mathbf{A}^g = \emptyset$ in the modified model). Finally, $Q(\{x_\infty\}|(x,T),a^i_\infty) = 1$ and $\lambda((x,T),a^g_\infty), Q(\cdot|(x,T),a^g_\infty)$ can be defined arbitrarily. Here $x_\infty$ is the absorbing cemetery with no future cost. It remains to put $C^{i,g}((x,t),a) = e^{\alpha t}C^{i,g}(x,t,a)$ for $t \in [0,T)$, where $C^{i,g}(x,t,a)$ are the original cost functions. If the $t$ component equals $T$ then $C^g((x,T),a^g_\infty)$ may be arbitrary and $e^{-\alpha T}C^i((x,T),a^i_\infty)$ can be equal to the terminal cost at the end of the planning horizon. After this rearrangement, the new (time-homogeneous, discounted) model is within the scope of the current work. Assumptions A,B,C will be satisfied if they are satisfied for the initial non-homogeneous model with horizon $[0,T]$.

In the similar way, one can take into account the time $\tau$ elapsed from the previous jump epoch: $\lambda(x,\tau,a)$ can describe the hazard rate of any general continuous distribution of the sojourn time, not necessarily exponential. After any one jump of the process, the new value of the $\tau$ component is zero.

According to Proposition 5.6, there is an optimal control strategy in the new model which is stationary and non-randomized. Of course, in the original setting, that strategy will be time-dependent.


## Acknowledgement.

This work was partially supported by JSPS KAKENHI Grant Number 26400215, by the French State, managed by the French National Research Agency (ANR) in the frame of the "Investments for the future" Programme IdEx Bordeaux - CPU (ANR-10-IDEX-03-02), RCMM and by EPSRC, Grant Ref: EP/M017206/1.


## References


[1] C.D. Aliprantis and K.C. Border. *Infinite dimensional analysis*. Springer, Berlin, third edition, 2006. A hitchhiker's guide.

[2] D.P. Bertsekas and S.E. Shreve. *Stochastic optimal control*, volume 139 of *Mathematics in Science and Engineering*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978. The discrete time case.

[3] O.L.V Costa and F. Dufour. *Continuous average control of piecewise deterministic Markov processes*. Springer Briefs in Mathematics. Springer, New York, 2013.

[4] O.L.V. Costa, F. Dufour, and A. Piunovskiy. Constrained and unconstrained optimal control of piecewise deterministic markov processes. *Submitted to SIAM Journal of Control and Optimization*, 2015.

[5] O.L.V. Costa and C.A.B. Raymundo. Impulse and continuous control of piecewise deterministic Markov processes. *Stochastics Stochastics Rep.*, 70(1-2):75–107, 2000.

[6] M.H.A. Davis. *Markov models and optimization*, volume 49 of *Monographs on Statistics and Applied Probability.* Chapman & Hall, London, 1993.

[7] M.A.H. Dempster and J.J. Ye. Impulse control of piecewise deterministic Markov processes. *Ann. Appl. Probab.*, 5(2):399–423, 1995.

[8] F. Dufour and A. Piunovskiy. Impulsive control for continuous-time Markov decision processes. *Adv. in Appl. Probab.*, 47(1):106–127, 2015.

[9] F. Dufour and A. Piunovskiy. Impulsive control for continuous-time markov decision processes: A linear programming approach. *To appear in Applied Mathematics and Optimization*, 2015.

[10] R.J. Elliott. *Stochastic calculus and applications*, volume 18 of *Applications of Mathematics (New York).* Springer-Verlag, New York, 1982.

[11] D. Goreac. Asymptotic control for a class of piecewise deterministic Markov processes associated to temperate viruses. *SIAM J. Control Optim.*, 53(4):1860–1891, 2015.

[12] O. Hernández-Lerma and J.-B. Lasserre. *Discrete-time Markov control processes*, volume 30 of *Applications of Mathematics (New York).* Springer-Verlag, New York, 1996. Basic optimality criteria.

[13] J. Jacod. *Calcul stochastique et problèmes de martingales*, volume 714 of *Lecture Notes in Mathematics.* Springer, Berlin, 1979.

[14] M. Schäl. On dynamic programming: compactness of the space of policies. *Stochastic Processes Appl.*, 3(4):345–364, 1975.

[15] A.A. Yushkevich. Verification theorems for Markov decision processes with controllable deterministic drift, gradual and impulse controls. *Teor. Veroyatnost. i Primenen.*, 34(3):528–551, 1989.